UNIVERSITÁ DEGLI STUDI DI ROMA “TOR VERGATA”

DOTTORATO DI RICERCA IN
TEORIA ECONOMICA E ISTITUZIONI
CICLO XXVII

Essays on latent approaches to economic growth and corruption

Alessandra Marcelletti

A.A. 2014/15

Tutor: prof. Giovanni Trovato
Coordinatore: prof. Luisa Corrado

Contents

Preface

This thesis studies the impact of unobserved socio - economic characteristics on the economy and social structure of different countries. To do so, we exploit and expand the specificity related to latent variable models. In particular, we focus on the possible interactions between latent traits and country-specific determinants of two phenomena: growth and corruption.
These statistical models simultaneously solve three empirical issues related to the nature of economic and social aspects: unobserved heterogeneity, omitted variables bias and non linearity in the regression function. In order to address these challenges, these models directly account for a latent variable into the estimation process. This identification strategy adjusts the estimation of the parameters of interest for unobserved heterogeneity due to the country's specific economic, political, and social environment.
The assumption underlying this thesis is that country - specific fundamentals could differently affect the dynamic of growth and corruption through unobservable and/or unmeasurable heterogeneous characteristics such as the quality of institutions, cultural attitude, subjective factors and so on. Thus, we relax the assumption that these fundamentals are invariant through different countries, allowing them to be grouped for the similarity of their features.
Entering into details, we analyze the dynamic of economic growth by applying a Finite Mixture Model and a Markov Switching Model with time varying transition probability. The final chapter provides a study of the dynamic of corruption by applying a Finite Mixture Model with concomitant variables.
The first paper explores the evolution and the variability of growth and per capita income by extending the empirical formulation of the augmented Solow model based on a multivariate-multidimensional specification. Following Bernanke's (2002) intuition that links investment in human and physical capital, population growth and long - run growth, we design an extension of the unidimensional finite mixture model based on the endogenous clustering approach lying on a bivariate multidimensional specification. Our estimation procedure solves for unobserved heterogeneity issues related to varying parameters across countries, omitted variables and non-linearities in the production function.
Our bivariate bidimensional discrete random effects model accounts for dependence between outcomes (i.e. per capita income and growth) and heterogeneity between countries. Furthermore, it models all parameters, not only the mean, as function of observable and unobservable covariates. In this way, we can investigate how economic factors affect the unpredictability of the economic performance, by explicitly modelling the variance of growth rate and per capita income.

As a by-product, the empirical evidence provides a posterior classification of countries sharing the same latent structure. In this respect, our contribution stands in highlighting strong heterogeneous characteristics within the groups of countries. In contrast with the existing literature we relax the assumption of per - capita Gross Domestic Product (GDP henceforth) and growth classified with the same posterior classification, and we take into account both observable and latent traits, analyzing the uncertainty and variability in the countries' economic performance.
The main idea of the second paper is that different growth paths are determined by switching from a growth regime to the other, i.e. miracle, stagnation, crisis and stability. The statistical model we apply is a Markov Switching Model with time varying transition probabilities. The empirical evidence is contained in a dataset on African Countries, where we study what is behind the asymmetries of output collapses, growth stability and different growth dynamics as a whole.
We believe that this identification strategy is the most appropriate, since Markov Switching models account for statistical issues such as non - linearity in the regression function, dependency weights and volatility in the dependent variable. We assume that the probability of switching from one growth path to the other depends on economic and political variables, allowing the switching process to be affected by latent traits. Hence, the growth rate dynamics are described as AR(1) process, while the state variable, that defines the different growth regime, is not completely latent but it depends on covariates assumed to be important determinants of the growth rate variation. Thus, our hypothesis is that switching from one political regime to the other, i.e. democracy, dictatorship and political instability, explains switching process of the growth rate, together with the observable variables, as openness to trade, foreign direct investment, exchange rate, and the unobservable country - specific characteristics specified in the estimation procedure.
The time - series of growth rate is therefore grouped into different regimes, accounting for the latent process and provide further evidence on the importance of specific aspects related to the institutional and socio - economic structure of each African country.
In the third paper we investigate the main drivers behind the literature's discordant results on the country determinants of corruption.
We identify them by applying a Finite Mixture Model with concomitant variables. This identification strategy fullfill the econometric challenge of omitting the unobservable heterogeneity between different countries. Our model includes this aspect in the estimation procedure, by imposing a latent structure for the covariates. We adopt the definition of corruption as a phenomenon that "occurs at the interface of the private and public sector" (Ackerman, 1997), whereas the decision of the agent in undertaking an activity that relates to corruption depends on the expected cost, the so called "risk of punishment" (Becker, 1968), the expected economic return and individual subjective factors included in the utility function. Nevertheless, despite latter elements, as tastes and preferences, propensity of committing illegal acts, sense of justice, as well as attitude towards risk, are powerful in explaining corruption, they are generally hidden and/or unmeasurable. Thus, to obtain a suitable model, we include them in our empirical

model as latent factors. Moreover, to partially adjust the estimation for the reverse causality between corruption and country-specific socio-economic structure, we estimate prior probabilities conditioning on initial measures of per capita GDP, fiscal rate and schooling.
As a result, the entire sample is divided into different groups, formed by country having the same socio-economic structure, given the observed (measurable) and unobserved (immeasurable) covariates.
Our empirical results provide evidence that there is strong heterogeneity among countries. To improve models about the dynamics of growth and corruption, we show that empirical formulation of these two phenomena should therefore take into account a latent variable in the estimation procedure, including presence of unobserved heterogeneity.
The thesis is structured as follows: the first chapter proposes a flexible bivariate finite mixture approach to understand cross-country differences in income per capita; the second chapter applies a Markov Switching Model with time varying transition probability to explore the stability and instability growth path in Sub-Saharan countries; the third chapter applies a Finite Mixture Model with concomitant variable to understand what drives the contradictory results in the literature regarding the determinants of corruption.

Chapter 1

A flexible bivariate location-scale finite mixture approach to economic growth

Alessandra Marcelletti*† Antonello Maruotti‡§, Giovanni Trovato¶

Abstract

We introduce a multivariate multidimensional mixed-effects regression model in a finite mixture framework. We relax the usual unidimensionality assumption on the random effects multivariate distribution. Thus, we introduce a multidimensional multivariate discrete distribution for the random terms, with a possibly different number of support points in each univariate profile, allowing for a full association structure. Our approach is motivated by the analysis of economic growth. Accordingly, we define an extended version of the augmented Solow model. Indeed, we allow all model parameters, and not only the mean, to vary according to a regression model. Moreover, we argue that countries do not follow the same growth process, and that a mixture-based approach can provide a natural framework for the detection of similar growth patterns. Our empirical findings provide evidence of heterogenous behaviors and suggest the need of a flexible approach to properly reflect the heterogeneity in the data. We further test the behavior of the proposed approach via a simulation study, considering several factors such as the number of observed units, times and levels of heterogeneity in the data.

Keywords: Country Classifications; Economic Growth; Finite Mixture Model.
JEL classification: O47, C14, C33, C39

1 Introduction

In modelling panel economic data, it is common to account for the unobserved heterogeneity between sample units, that is, the heterogeneity that cannot be explained by means of observable covariates (see e.g. Wooldridge, 2002 ; Fitzmaurice et al., 2008). This is normally accomplished by the introduction of latent variables or random effects. For instance, a typical approach consists of associating a random intercept to every sample unit which affects the distribution of each time-specific response in the same fashion. This allows us to account for a form of unobserved heterogeneity which is due to unobservable covariates and related factors. The above considerations are obviously pertinent when we deal with economic growth

*Dipartimento di Economia Diritto ed Istituzioni, Università di Roma "Tor Vergata" via Columbia 2 00133 Roma - Italy, e-mail: alessandra.marcelletti@uniroma2.it

†Dipartimento di Scienze Politiche, Luiss Guido Carli, Viale Romania 32, 00197 Roma, email: amarcelletti@luiss.it

‡Southampton Statistical Sciences Research Institute, University of Southampton, SO17 1BJ, UK, email: a.maruotti@soton.ac.uk

§Dipartimento di Scienze Politiche, Università di Roma Tre, Via Gabirello Chiabrera 199, 00145, Roma - Italy, email: antonello.maruotti@uniroma3.it

¶Dipartimento di Economia Diritto ed Istituzioni, Università di Roma "Tor Vergata", via Columbia 2 00133 Roma - Italy, tel:+39-0672595649

modelling, where sample units (i.e. countries) are characterized by heterogeneous income performances. Addressing the heterogeneity of analyzed processes is of fundamental importance to the study to the economic growth and has led to a substantial evidence for the existence of variations in growth patterns across countries. Indeed, since Solow's seminal paper (1956), different econometric and statistical approaches are used to look at countries' growth. Dynamic panel data with fixed effect (Caselli et al., 1996; Islam, 1995; Temple, 1999), as well as extreme bound analysis (Levine and Renelt, 1992; Temple, 2000), Bayesian model averaging (Doppelhofer et al., 2000; Fernandez et al., 2001) or model on varying coefficients are performed to deal with the main empirical challenges in growth theory: unobserved heterogeneity (Caselli et al., 1996; Pesaran and Smith, 1995; Lee et al., 1997; Durlauf and Johnson, 1995), uncertainty (Temple, 2000) and omitted variable bias (Durlauf and Quah, 1999).

Recently, data-driven approaches to estimate multiple (heterogeneous) growth processes have been employed within the wide class of mixture models (Alfó et al., 2008, Owen et al., 2009; Kerekes, 2012; Baştürk et al., 2012; Bertarelli and Bernardini Papalia, 2013).

We propose an approach to panel growth data based on a flexible bivariate location-scale finite mixture approach, which may be seen as an extension of the approach introduced by Alfó et al. (2008). We introduce a bivariate bidimensional discrete random effects model to account for dependence between outcomes (i.e. per capita income and growth) and heterogeneity between countries in the augmented Solow growth model. The proposed approach may be cast in the literature about finite mixture models for panel data. It is worth noting that other extensions of the finite mixture approach for panel data are available in the literature. We mention, in particular, the extensions proposed by Pittau et al. (2010) and Martínez-Zarzoso and Maruotti (2011), where countries are clustered into *clubs* depending on unobserved characteristics. Moreover, our approach is more general than those of Durlauf and Johnson (1995) and Ardıç (2006) in which clustering is performed beforehand (i.e. clustering is exogenously specified). Indeed, we develop an endogenous clustering approach lying on a bivariate bidimensional model recovering Bernanke and Gürkaynak (2002) intuition: country's rate of investment and of human capital and the population growth rate are correlated with its long run growth of output per capita. Thus we contribute to this branch of literature by providing an empirical formulation of the augmented Solow model based on a multivariate-multidimensional specification, that allows to solve the unobserved heterogeneity issue. We address the heterogeneity issues related to: varying parameters across countries, omitted variables and non-linearities in the production function. Indeed, the incorrect specification of the country-specific effects leads to inconsistent parameter estimation, generating omitted variable bias (Caselli et al., 1996).

As a by-product, we provide a posterior classification of countries sharing the same latent structure, highlighting strong heterogeneous behaviours. With respect to the existing approaches, we relax the assumption of the same posterior classification for the gross domestic product (GDP) per capita level and the growth rate. This allows us to let free the posterior classification given the observed variable and the latent effect, and to analyze the uncertainty and the variation in the different economics performance. We are able to distinguish between *between* group, and *within* group variations allowing for the human and physical capital and the population growth rate to simultaneously affect the different country growth experience, in terms of growth path and variability in the GDP per capita and growth rate. We further allow for explicitly modelling the scale parameter as a function of covariates. Indeed, we introduce two separates equations for the location and scale parameters of the dependent variables, such that the explanatory variables are associated not only to high or low values of the dependent variable, but also to the unpredictability of the variable itself.

Computational complexity is often the price we have to pay to flexibility. However, we show that parameter estimates can be obtained by extending the Expectation-Maximization (EM) algorithm (Dempster et al., 1977) for finite mixture to the multidimensional case. Furthermore, we avoid any restriction on the covariance structure of the random effects as assumed e.g. by the so-called one-factor model (Winkelmann, 2000), which is more parsimonious but could be hard to justify in empirical applications. By allowing the number of mixture components to grow with the sample size, the proposed model can be also used as a semiparametric estimator of multivariate mixed effects models, where the distribution of the random effects is estimated by a discrete multivariate random variable with a finite number of support points. This can be seen as a possible solution to computational issues arising with multivariate mixed models.

We illustrate the proposal by a simulation study in order to investigate the empirical behaviour of the proposed approach with respect to several factors, such as the number of observed units and times and the distribution of the random term (with varying number of support points). Finally, we test the proposal by analysing a sample taken from the Summers-Heston Penn World Tables (PWT) version 8.0 for years 1975-2005 for non-oil countries. We identify a set of variables that affect the volatility of economic growth and remark the importance of including *baseline* GDP as a covariate in the model specification. Moreover, different levels of heterogeneity are detected in GDP and GDP growth, respectively. More precisely, we find that our sample is much more heterogeneous with respect to GDP levels than growth patterns. Although this result sounds obvious, previous empirical results, based on unidimensional specification of the latent structure, were not able to distinguish for different heterogeneity levels (see e.g. Alfó et al., 2008). Instead, our approach can easily accommodate for different heterogeneity levels in the univariate profiles and, simultaneously, accounts for association between outcomes. About obtained results, we get two clusters representing high-growth and low-growth countries, and six clusters are identified with respect to GDP levels.

The plan of the paper is as follows. In Section 2, we specify the proposed model in a general form and in Section 3 we provide the computational aspects of the adopted maximum likelihood algorithm. In Section 4, we give a comparison of the performance of several model specifications under different data generation schemes by means of a simulation study. In Section 5, we present an empirical application on real world data motivating this paper. In Section 6, we point out some remarks, along with drawbacks that may arise by adopting the proposed methodology.

2 Statistical framework

We start assuming that the analysed sample is composed of n statistical units (e.g. countries): continuous responses y_{itj}, corresponding to $(j = 1, \ldots, J)$ outcomes and two vectors of covariates $\mathbf{x}'_{itj} = (1, x_{itj1}, \ldots, x_{itjP_j})$ and $\mathbf{z}'_{itj} = (1, z_{itj1}, \ldots, z_{itjQ_j})$, which can vary over outcomes, are recorded for each unit i $(i = 1, 2, \ldots, n)$ at time t $(t = 1, 2, \ldots, T)$. Following the usual notation for longitudinal multivariate data, let $\mathbf{y}_{it} = (y_{i11}, \ldots, y_{itJ})'$ denote the vector of observed responses for unit i at the t-th time. We assume that y_{itj} are realizations of conditionally independent random variables, with parameters $\boldsymbol{\theta}_{itj} = (\theta_{itj1}, \theta_{itj2}, \ldots, \theta_{itjM})$. When we face multivariate analysis, and the primary focus of the analysis is not only to build a regression model, but even to describe association among responses, the univariate approach is no longer sufficient and needs to be extended. In this context, we are likely to face complex phenomena which can be characterized by having a non-trivial correlation structure. For instance, omitted covariates may affect more than one response; hence, modelling the association among the outcomes can be a fundamental aspect of research. Beyond that, the association structure could be of interest by itself, as we may be interested in understanding the nature of the stochastic dependence among the analysed phenomena. Furthermore, it is well known that, when responses are correlated, the univariate approach is less efficient than the multivariate one, since in estimating the parameters in the single equations, the multivariate approach takes into account of zero restrictions on parameters occurring in other equations (for a detailed discussion on this topic see e.g. Zellner, 1962; Davidson and MacKinnon, 1993).

A standard way to insert dependence among responses is to assume that they share some common latent structure. Thus, the model specification is completed by connecting the J univariate submodels through a common latent structure, represented by a set of random effects $\mathbf{u}_i = (\mathbf{u}_{i1}, \ldots, \mathbf{u}_{iJ})$ which account for potential heterogeneity among statistical units and correlation between outcomes. In a regression setting, the interest is usually focused upon the mean which is modelled through a linear mixed model, providing a very broad framework for modelling dependence in the data (Verbeke et al., 2014). Nevertheless, statistical models rarely allow the modelling of parameters other than the mean of the response variable as functions of the explanatory variables. For instance, the scale parameter is usually not modelled explicitly in terms of the explanatory variables but implicitly through its dependence on the mean. In the following, we relax such a constrain and define a location-scale multivariate regression framework by specifying J conditionally independent (given the covariates and the random effects) regression models. Let us decompose the design vector as $\mathbf{x}_{itj} = \{\mathbf{x}_{itj}^{(1)}, \mathbf{x}_{itj}^{(2)}\}$, where the variables whose effects are assumed to be fixed are collected in $\mathbf{x}_{itj}^{(1)}$, while those which vary across units are in $\mathbf{x}_{itj}^{(2)}$. The M-dimensional parameter vector $\boldsymbol{\theta}_{itj}$ is related to covariates and random effects. Let us specify θ_{itj1} as the location parameter, θ_{itj2} as the scale parameter and θ_{itj3} as a shape parameter (whenever needed) and let $g_m(\cdot)$ be a known monotonic link function relating $\theta_{itjm}, m = 1, \ldots, 3$ to covariates and random effects, we define the following regression models

$$\begin{cases} g_1(\theta_{itj1}) = \mathbf{x}_{itj}^{\prime(1)} \boldsymbol{\lambda}_j + \mathbf{x}_{itj}^{\prime(2)} \mathbf{u}_{ij} \\ \\ g_2(\theta_{itj2}) = \mathbf{z}_{itj}^{\prime} \boldsymbol{\gamma}_j \\ \\ g_3(\theta_{itj3}) = \tilde{\boldsymbol{\gamma}}_j \end{cases} \tag{1}$$

where $\mathbf{u}_{ij}$ represents unit- and outcome-specific random effects, drawn from a multivariate parametric density, $\boldsymbol{\lambda}_j$, $\boldsymbol{\gamma}_j$ and $\tilde{\boldsymbol{\gamma}}_j$ are outcome- and moment-specific fixed parameters. Of course, covariates may be included in the shape-parameter model, but this may complicate results interpretation in empirical applications.

Given the model assumptions, the likelihood function can be written as follows:

$$L(\cdot) = \prod_{i=1}^{n} \left\{ \int_{\mathcal{U}} \prod_{j=1}^{J} \prod_{t=1}^{T} f(y_{itj} \mid \mathbf{u}_{ij}, \mathbf{x}_{itj}, \mathbf{z}_{itj}) b(\mathbf{u}_i) d\mathbf{u}_i \right\} \tag{2}$$

where $f(\cdot)$ is a generic probability density function, $\mathcal{U}$ represents the support for $b(\mathbf{u}_i)$, the distribution function of $\mathbf{u}_i$, with $E(\mathbf{u}_i) = 0$.

Although, at first glance, the approach proposed so far is appealing, it has several computational drawbacks and limitations. Indeed, the random effects distribution is unknown and assuming a multivariate Gaussian distribution may be a too strong and unverifiable assumption

and, moreover, may affect parameters estimate. Indeed, in some situations, the distribution of the random effects may depart from normality. This problem has been addressed, for example, by specifying a different parametric distribution family for the random terms, such as multivariate skewed and/or heavy-tailed distributions (Ferreira and Steel, 2006; (Ferreira and Steel, 2004)). An alternative approach is to use nonparametric maximum likelihood based on finite mixtures, which provide a more flexible framework to deal with departure from normality of the random effects distribution (see e.g. Böhning, 1995; Aitkin, 1999). Nevertheless, even if the latter is computationally efficient when compared to parametric random effect models, it is intrinsically unidimensional, since it is based on a single categorical latent variable. This may lead to problems when the task is testing for dependence between the random effects. Indeed, the model under independence does not occur as a special case of the dependence model.
In the following, we consider a J-variate J-dimensional latent structure such that the independence model is nested in the multivariate one, and different levels of heteorgeneity in the J univariate profiles can be identified. In order to specify a latent structure of this kind, we leave the distribution of the random effect $b(\cdot)$ completely unspecified and invoke the non-parametric maximum likelihood approach.

Formally, random effects distribution can be approximated through a discrete distribution with $K_j \leq n$ support points at the marginal level. Mass joint probability $\pi_{k_1,k_2,\ldots,k_J}$ are attached to location $(\mathbf{u}_{k_1}, \mathbf{u}_{k_2}, \ldots, \mathbf{u}_{k_J})$ for $k_j = 1, \ldots, K_j$. Focusing on the bivariate ($J = 2$) case, without lacking of generality, we define the following location-scale multivariate regression model

$$\begin{cases} g_1(\theta_{itj1}) = \mathbf{x}_{itj}^{'(1)} \boldsymbol{\lambda}_j + \mathbf{x}_{itj}^{'(2)} \mathbf{u}_{k_j} \\ \\ g_2(\theta_{itj2}) = \mathbf{z}'_{itj} \boldsymbol{\gamma}_j \\ \\ g_3(\theta_{itj3}) = \tilde{\boldsymbol{\gamma}}_j \end{cases} \tag{3}$$

According to model assumptions, the likelihood function in the bivariate case is given by

$$L(\cdot) = \prod_{i=1}^{n} \left\{ \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \pi_{k_1 k_2} \prod_{j=1}^{2} \prod_{t=1}^{T} f(y_{itj} | \mathbf{x}_{itj}, \mathbf{z}_{itj}, \mathbf{u}_{i1} = \mathbf{u}_{k_1}, \mathbf{u}_{i2} = \mathbf{u}_{k_2}) \right\} \tag{4}$$

where $\pi_{k_1 k_2} = Pr(\mathbf{u}_{i1} = \mathbf{u}_{k_1}, \mathbf{u}_{i2} = \mathbf{u}_{k_2})$ is the joint probability associated to each couple of locations $(\mathbf{u}_{k_1}, \mathbf{u}_{k_2})$. The following constraints hold $\sum_{k_1=1}^{K_1} \pi_{k_1} = \sum_{k_2}^{K_2} \pi_{k_2} = \sum_{k_1 k_2} \pi_{k_1 k_2} = 1$

with

$$\pi_{k_1} = Pr(\mathbf{u}_{i1} = \mathbf{u}_{k_1}) = \sum_{k_2=1}^{K_2} \pi_{k_1 k_2}$$

and

$$\pi_{k_2} = Pr(\mathbf{u}_{i2} = \mathbf{u}_{k_2}) = \sum_{k_1=1}^{K_1} \pi_{k_1 k_2}.$$

We would remark that the number of locations (i.e. mixture components) may vary between outcomes. Thus, we control for heterogeneity in the univariate profiles and for the association between latent effects in the two profiles. This approach results in a finite mixture with $K_1 \times K_2$

components, in which each of the K_1 locations are coupled with each of the K_2 locations of the second outcome. If $J = 1$, our proposal reduces to a univariate finite mixture model.

3 Computational details

Let $\tilde{\boldsymbol{\theta}}$ be a short-hand notation for all non-redundant models parameters corresponding to the vectors $(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \tilde{\boldsymbol{\gamma}}, \boldsymbol{\pi}, \boldsymbol{u})$, inference for the proposed model is based on log-transformation of the likelihood in (4).

To estimate $\tilde{\boldsymbol{\theta}}$, we maximized the log-transformation of (4) by using a version of the EM algorithm (Dempster et al., 1977). The EM algorithm alternates the following steps until convergence

E-step: compute the conditional expected value of the complete data log-likelihood given the observed data and the current estimate of model parameters; and

M-step: maximize the preceding expected value with respect to $\tilde{\boldsymbol{\theta}}$.

Let $w_{ik_1k_2}$ denote a dummy variable equal to 1 if unit i is in component k_1 and k_2 in the two univariate profiles, respectively, and zero otherwise. The complete data likelihood, which we would compute if we knew these dummy variables, is

$$L_c(\cdot) = \prod_{i=1}^{n} \left[\sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \pi_{k_1k_2} f_{ik_1k_2} \right]^{w_{ik_1k_2}} \tag{5}$$

And its corresponding log-transformation is

$$\ell_c(\cdot) = \sum_{i=1}^{n} \sum_{k_1}^{K_1} \sum_{k_2}^{K_2} w_{k_1k_2} \left\{ \log(\pi_{k_1k_2}) + \log f_{ik_1k_2} \right\} \tag{6}$$

where $f_{ik_1k_2} = f_{ik_1} f_{ik_2} = \prod_{t=1}^{T} f(y_{it1}|\mathbf{x}_{it1}, \mathbf{z}_{it1}, u_{k_1}) f(y_{it2} \mid \mathbf{x}_{it2}, \mathbf{z}_{it2}, u_{k_2})$.

The conditional expected value of $\ell_c(\cdot)$ at the E-step has then the same expression as given previously in which we substitute the variable $w_{ik_1k_2}$ with its corresponding expected value

$$\hat{w}_{k_1k_2} = \frac{\pi_{k_1k_2} f_{ik_1k_2}}{\sum\limits_{k_1k_2} \pi_{k_1k_2} f_{ik_1k_2}}. \tag{7}$$

where $\hat{w}_{k_1k_2}$ is the posterior probability the the i-th unit belongs jointly to the k_1 and k_2 components of the mixture. We can easily get the marginal posterior probabilities

$$\hat{w}_{ik_1} = \sum_{k_2} \hat{w}_{ik_1k_2} \quad \hat{w}_{ik_2} = \sum_{k_1} \hat{w}_{ik_1k_2} \tag{8}$$

At the M-step, the conditional expected value of (6) is maximized by separately maximizing its components. Indeed, the score function is

$$\sum_{i=1}^{n} \sum_{k_1}^{K_1} \sum_{k_2}^{K_2} w_{k_1k_2} \frac{\partial}{\partial \boldsymbol{\theta}} \left\{ \log(\pi_{k_1k_2}) + \log f_{ik_1} + \log f_{ik_2} \right\}.$$

Let us partition the parameter vector $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_{k_1}, \tilde{\boldsymbol{\theta}}_{k_2})$, where $\tilde{\boldsymbol{\theta}}_{k_j}$ collects the parameters of the j-th profile such that

$$\frac{\partial \ell(\cdot)}{\partial \tilde{\boldsymbol{\theta}}_{k_1}} = \sum_{i=1}^{n} \hat{w}_{ik_1} \frac{\partial}{\partial \tilde{\boldsymbol{\theta}}_{k_1}} \log(f_{ik_1}); \tag{9}$$

$$\frac{\partial \ell(\cdot)}{\partial \tilde{\boldsymbol{\theta}}_{k_2}} = \sum_{i=1}^{n} \hat{w}_{ik_2} \frac{\partial}{\partial \tilde{\boldsymbol{\theta}}_{k_2}} \log(f_{ik_2}) \tag{10}$$

and

$$\frac{\partial \ell(\cdot)}{\partial \pi_{k_1 k_2}} = \sum_{i=1}^{n} \hat{w}_{ik_1k_2} \frac{\partial}{\partial \pi_{k_1 k_2}} \log \pi_{k_1 k_2} \tag{11}$$

An explicit solution is available to maximize the last M-step equation, which consists of

$$\hat{\pi}_{k_1 k_2} = \frac{\sum_{i=1}^{n} \hat{w}_{ik_1k_2}}{n}.$$

To maximize the other two parts, we can use a standard iterative algorithm of Newton-Raphson type for linear mixed models. We take the value of $\tilde{\boldsymbol{\theta}}$ at convergence of the EM algorithm as the maximum likelihood estimate. As it is typical for finite mixture models the likelihood may be multimodal and the point at convergence depends on the starting values for the parameters, which then need to be carefully chosen. In this regard, we run the EM algorithm from multiple random starting points for a number of steps, then pick the one with the highest likelihood, and continue the EM from the picked point until convergence. However, other methods can be used; for example, a gradient function based on directional derivatives can be used to get optimality criteria (see e.g. Wang, 2010).

At last, we approach the model selection problem by looking at penalized likelihood criteria, Akaike information criterion (AIC) and Bayesian information criterion (BIC). In this way we select the number of mixture components and we can also compare the different models. BIC, achieved in the Bayesian framework is found to be satisfactory in the model-based clustering context (see among others Fraley and Raftery, 2002, for further details). Both criteria are likelihood based and they differ for the different penalization used. In fact, denoting with d the number of independent parameters to be estimated and with n the sample size, BIC is obtained as $BIC = -2\ell(.) + d\ln(n)$, and AIC is given by $AIC = -2\ell(.) + 2 * d$.

4 Simulation study

To assess the properties of the maximum likelihood estimator described in Section 3, we carried out a simulation study, which is described subsequently. The same study allows us to assess the goodness of classification.

4.1 Simulation design

We considered two scenarios: the first with two response variables (both Gaussian-distributed) with $K_1 = K_2 = 2$ mixture components each and the second with higher heterogeneity levels, i.e. by defining a bivariate model with $K_1 = 2$ and $K_2 = 3$ mixture components for each outcome respectively. Under each scenario, we considered two continuous covariates, one in the linear predictor for the mean and one in the regression model for the scale parameter, and generated 500 samples from the proposed model with $T = 5; 10$ (panel length) and $n = 100; 1000$ (sample size). Under this setting, $\boldsymbol{\theta}_{itj} = (\theta_{itj1}, \theta_{itj2}) = (\mu_{itj}, \sigma_{itj})$

Scenario 1. We assume that the outcomes are conditionally independent and proceeded to generate 500 samples from

$$Y_{it1} \mid \mu_{it1}, \sigma_{it1} \sim N(\mu_{it1}, \sigma_{it1})$$

$$Y_{it2} \mid \mu_{it2}, \sigma_{it2} \sim N(\mu_{it2}, \sigma_{it2})$$

where the following bivariate regression model (with a single covariate) holds

$$\mu_{it1} = u_{k_1} + \lambda_{11} x_{it} = \begin{cases} -1 + 0.5x_{it}, & k_1 = 1 \\ 1 + 0.5x_{it}, & k_1 = 2 \end{cases}$$

$$\log(\sigma_{it1}) = \gamma_{01} + \gamma_{11} z_{it} = 0.5 + 0.75 z_{it}$$

and

$$\mu_{it2} = u_{k_2} + \lambda_{12} x_{it} = \begin{cases} 2 + 0.5x_{it}, & k_2 = 1 \\ -2 + 0.5x_{it}, & k_2 = 2 \end{cases}$$

$$\log(\sigma_{it2}) = \gamma_{02} + \gamma_{12} z_{it} = 1 + 0.25 z_{it}$$

with

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} = \begin{bmatrix} 0.4 & 0.1 \\ 0.2 & 0.3 \end{bmatrix}$$

Scenario 2. We assume that the outcomes are conditionally independent and proceeded to generate 500 samples from

$$Y_{it1} \mid \mu_{it1}, \sigma_{it1} \sim N(\mu_{it1}, \sigma_{it1})$$

$$Y_{it2} \mid \mu_{it2}, \sigma_{it2} \sim N(\mu_{it2}, \sigma_{it2})$$

where the following bivariate regression model (with a single covariate) holds

$$\mu_{it1} = u_{k_1} + \lambda_{11} x_{it} = \begin{cases} -1 + 0.5x_{it}, & k_1 = 1 \\ 1 + 0.5x_{it}, & k_1 = 2 \end{cases}$$

$$\log(\sigma_{it1}) = \gamma_{01} + \gamma_{11} z_{it} = 0.5 + 0.75 z_{it}$$

and

$$\mu_{it2} = u_{k_2} + \lambda_{11} x_{it} = \begin{cases} 2 + 0.5x_{it}, & k_2 = 1 \\ -2 + 0.5x_{it}, & k_2 = 2 \\ 0 + 0.5x_{it}, & k_2 = 3 \end{cases}$$

$$\log(\sigma_{it2}) = \gamma_{02} + \gamma_{12} z_{it} = 1 + 0.25 z_{it}$$

with

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \end{bmatrix} = \begin{bmatrix} 0.1 & 0.1 & 0.2 \\ 0.2 & 0.3 & 0.1 \end{bmatrix}.$$

4.2 Simulation results

For each sample, we computed the maximum likelihood estimate of the parameters and the corresponding standard errors, under the assumed model. We also evaluate the performance of the proposed in correctly clustering the statistical units into mixture components. The Rand Index (Hubert and Arabie, 1985) is considered. The true matrix $\mathbf{W} = \{w_{ik_1k_2}\}$ of component membership and the crispy estimated matrix $\mathbf{W}^* = \{w^*_{ik_1k_2}\}$, where each element $w^*_{uk_1k_2}$ is defines as

$$w^*_{uk_1k_2} = \begin{cases} 1 & \text{if} k_1, k_2 = \arg\max_{k_1,k_2} \hat{w}_{ik_1k_2} \\ 0 & \text{otherwise} \end{cases}$$

are compared. Formally, let $n_{k_1k_2}$ denote the number of all pairs of data points which are either put into the same cluster by both partitions or put into different clusters by both partitions. Conversely, let $n^*_{k_1k_2}$ denote the number of all pairs of data points that are put into one cluster in one partition, but into different clusters by the other partition. The partitions disagree for all pairs $n^*_{k_1k_2}$ and agree for all pairs $n_{k_1k_2}$. We can measure the agreement by the Rand index $n_{k_1k_2}/(n_{k_1k_2} + n^*_{k_1k_2})$ which is invariant with respect to permutations of cluster labels.

For *Scenario 1*, the simulation results in terms of bias and standard deviation of the maximum likelihood estimator of each parameter of interest are shown in Table 1, together with the Rand Index. We can observe that, the bias of each estimator is always low and decreases as T increase; moreover, its standard deviation decreases. Indeed, for $n = 100$ and $T = 10$ the estimators are unbiased. By increasing the number of available times, the clustering performance improves as well as shown by the Rand Index. For sake of brevity, we do not report the results for $n = 1000$. They do not provide any further insight to the already discussed results.

By considering *Scenario 2*, in which a higher degree of heterogeneity is assumed in one of the two outcomes, we can easily detect a different estimators behavior (see Table 2). Obviously, for small sample size ($n = 100$) and $T = 5$, higher bias and standard deviations are estimated with respect to those in *Scenario 1*. However, estimates variability decreases at the expected rate of $\sqrt{n}$ with respect to n and at a faster rate with respect to T. By increasing the sample size to $n = 1000$, we get less biased estimates, as expected. Clustering performances are sensitive to n and T as well. Indeed ,the larger is the sample size the better is the recovered latent structure.

5 Empirical framework

5.1 Data

The sample is composed by an unbalanced panel of 101 countries over the period 1975-2010. Data on the dependent variables and the investment share on physical capital *(sk)* are retrieved from the Heston-Summers-Aten dataset (Penn World Table 8.0). Data on human capital *(sk)*, measured as the total enrollment in secondary education, is retrieved from the World Bank. From the same database, we also collect: openness to trade *(open)*, measured as the sum of exports and imports as share of GDP, and the credit to the Private Sector as a fraction of GDP *(fin)*, used as a proxy for financial development. In order to understand the effect of financial factor on the growth fluctuations through the household consumption channel, the private sector on GDP is preferred as measure since it does not account for the credit provided from the Central and development bank to the public sector. Government consumption *(govcons)* is calculated as the general government final consumption expenditure (as share of GDP).

Unemployment rate*(unempl)* and the inflation level (*infl*) are obtained from the Penn World Table 8.0 dataset.

In order to avoid the endogeneity problems related to growth model estimation, we consider non-overlapping 5-year period with explanatory variable averaged over the corresponding time period; while the dependent variables are taken 5 periods ahead (Bond et al., 2001). Indeed, endogeneity could be due to the fact that "country-specific heterogeneity cannot be captured if one does not look at between-countries variation which cannot be explained by observed covariates but remains persistent over the analysed time period." (Alfó et al., 2008, pg. 495). Thus, the dependent variables are the average of GDP per capita over the 5-years period (y_{it1}), and the average annual growth of real GDP over the same non overlapping period (y_{it2}). Table 3 provides descriptive statistics, variables description, and data sources.

To analyze the marginal distribution of the response variables, graphical and statistical analysis are provided. Figure 1 displays a clear multimodal distribution for the GDP level, supporting the idea of different sub-populations in the outcome. The marginal distribution of growth rates does not show any multimodality, although a small bump can be detected on the left with respect to the distribution mode. However, we cast some doubts that growth rate follows a Gaussian distribution. Thus, to complement the graphical analysis, Shapiro-Wilk and Jarque-Bera tests and summary statistics are provided in Table 4 for the two outcomes. Skewness and kurtosis of each response variable indicate a departure from the normal distribution. Whilst, it is expected that both Shapiro-Wilk and Jarque-Bera tests indicate departure from marginal normality for the GDP level, we obtain a significant departure from normality for the growth rate outcome as well. Thus, we opt for a (mixture of) heavy-tailed distribution to properly model growth rates.

5.2 Economic growth

To understand the cross-country differences in income performances and to account for dependence between per capita income and growth, we introduce a flexible bivariate multidimensional finite mixture approach for the location and the scale parameters, and for the shape parameter when it is required, as described in Section 2. To jointly determine the evolution of income per capita and volatility of growth, instead of modelling the scale parameter through the dependence on the mean, we explicit the variance of the growth rate as dependent on explanatory variables. Thus, growth determinants are associated not only to high or low values of the dependent variable but also to unpredictability of the variable itself.

Formally, for each country i at time t, let the GDP level (y_{it1}) be a Gaussian random variable, i.e. $y_{it1} \sim N(\mu_{it1}, \sigma_{it1})$, and the GDP growth rate (y_{it2}) be t-distributed to account for heavy tails in the growth distribution, i.e. $y_{it2} \sim t(\mu_{it2}, \sigma_{it2}, \nu_{it2})$. To explore the determinants of both growth level and growth volatility, we choose variables found to be robust in the economic growth literature (see e.g. Levine and Renelt, 1992; Mankiw et al., 1992; Cecchetti et al., 2006), and define the following mixed-effects regression model for y_{it1}

$$\begin{cases} \mu_{it1} = u_{i10} + \lambda_{11} sk_{it} + \lambda_{21} sh_{it} + \lambda_{31}(n_{it} g\delta) \\ \\ \log(\sigma_{it1}) = \gamma_{01} \end{cases} \tag{12}$$

where sk_{it} and sh_{it} are the share of output invested in physical and human capital, respectively, δ is the depreciation rate, n is the population growth rate and g is the technological progress. As it is common in the growth literature, the term $g + \delta$ is assumed to be common across countries and equal to 0.5. Parameters in model (12) capture the effect of the human and physical capital accumulation process, and the population growth on the income per capita. They can be explicit as:

$$\lambda_{11} = \frac{\alpha}{(1-\alpha-\beta)} \quad \lambda_{21} = \frac{\beta}{(1-\alpha-\beta)} \quad \lambda_{31} = \frac{\alpha+\beta}{(1-\alpha-\beta)} \tag{13}$$

where α and β are respectively the share of physical and human capital, such that $(\alpha+\beta) < 1$. It is worth noting that the λ_{11} and λ_{21} are expected to be positive, while λ_{31} to be negative, since human and physical capital accumulation boost economic growth, while the population growth rate is thought to discourage the evolution of the economy (see among others Solow, 1956; Mankiw et al., 1992; Barro, 1991). The random intercept u_{i10} is let free to vary across countries since it captures the unobserved heterogeneity due to the omission and/or the immeasurable nature of some country-specif factors.

According to Bernanke and Gürkaynak (2002), the definition of the augmented Solow model implies a bivariate growth model, in which the long run growth of output per capita is correlated with the accumulation of human and physical capital and the population growth rate. We adopt a reduced-form model for the location parameter of the growth rate (see Goetz and Hu, 1996for further details) such that

$$\begin{cases} \mu_{it2} = u_{i20} + u_{i21}\ln(yc_{it}) \\ \log(\sigma_{it2}) = \gamma_0 + \gamma_{12} unempl_{it} + \gamma_{22} fin_{it} + \gamma_{32} infl_{it} + \gamma_{42} open_{it} + \gamma_{52} govcons_{it} \\ \nu_{it2} = \tilde{\gamma}_{02} \end{cases} \tag{14}$$

The random coefficient u_{i21} (attached to the initial level of income per capita) controls for the transitional dynamics affecting the evolution of the growth rate. It is worth recalling that the neoclassical approach predict a fixed and negative coefficient for the initial level of income per capita $lnyc_{it}$ accounting for country convergence.

In our approach, economic stability is directly modelled by including an equation for the variance of the growth rate, that regress the unpredicatability of the response variable on financial development, international openness, government consumption, inflation and unemployment rate (e.g., Cecchetti et al., 2006; Giovanni and Levchenko, 2009). We expect that cyclical variables (unemployment rate and inflation) have a destabilizing effect on growth, i.e. γ_{12} and γ_{32} are expected to be positive, while financial development and government consumption decrease growth volatility. The effect of openness to trade on economic growth is still debated in the literature.

Again, the random terms u_{i02} and u_{i21} in the location parameter's equation are let free to vary among countries and response variables, by allowing for a random slope as well. This allows us to simultaneously understand the variation across country in the standard of living and in the volatility of the outcome per capita, leaving the posterior classification of the mixture model to be free to vary among outcomes.

5.3 Results

A major research question would concern the need of a complex model like the one we introduce to properly model economic growth. Thus, to remark the crucial role of the bivariate approach with respect to the univariate one, we start our empirical analysis by comparing univariate and multivariate approaches. Firstly, we fit univariate mixed-effects models for each outcome separately, with $K_1 = 2, \dots, 7$ and $K_2 = 2, \dots, 4$. Model selection results are provided in Table 5, and models with $K_1 = 6$ and $K_2 = 2$, respectively, are selected. Similarly, we perform model selection for the bivariate model specified in the previous section, with varying $K_1 = 2, \dots, 7$ and $K_2 = 2, \dots, 4$. In the bivariate case the AIC is in favour of the $K_1 = 6$ and $K_2 = 3$, while the BIC select the model with $K_1 = 6$ and $K_2 = 2$ groups (see Table 6). By comparing penalized likelihood criteria, it is clear that linking the two univariate profiles by a shared (correlated) random effects structure, i.e. adopting a bivariate approach, leads to better results in terms of trade-off between model fit and model complexity. In the following we look at the results obtained with the bivariate selects according to the BIC. This choice is motivated by looking at parsimony and for comparison purposes (with respect to univariate model specifications). In Figure2 we provide evidence of the goodness of fit of the proposed model, and of the relatively small increase in goodness of fit the $K_1 = 6$ and $K_2 = 3$ model selected according to the AIC. The Parameter estimates are provided in Table 7. The main difference between the univariate and the multivariate approaches is on the magnitude of covariates effects in the equation for the mean of GDP level. Indeed, the bivariate approach parameter estimates confirm the augmented Solow model intuition, i.e. the accumulation process of physical and human capital exhibits more reasonable value of the coefficients with respect to univariate case. As discussed before, the intercept term captures the omitted country-specific features, such as, above all, institutional characteristic. This is related to the idea that accumulation driven growth equation is incomplete (see e.g. Alfó et al., 2008), and, coherently with the literature, the highest value for the random effect is found for the component clustering the richest and more industrialized countries, such as USA and UK. However, we will investigate the obtained clustering in depth in Section 5.4.

As formalized before, the location parameter for the growth rate is estimated by applying a reduced-form model where the independent variables is the 5-years backward value of GDP per capita. This allows for avoiding biased estimation in the parameters due to the dependence among physical and human capital on income per capita (Goetz and Hu, 1996). Furthermore, to account for the difference in initial level of GDP per capita, we leave the initial level of GDP to vary among countries. Results show the existence of two groups: the first group characterized by a negative and significant effect of the initial level of GDP on the growth pattern, confirming economics theory about convergence; the second group is characterized by the possible existence of multipla equilibria and the lack of convergence. These results suggest the presence of a convergence club, that is, a group of countries with different levels of per capita real GDP within which countries converge to a group-specific growth path, i.e. the neoclassical prediction of the convergences is proved for those countries. The second component, clustering low income countries, shows lack of income convergence allowing for the potential existence of multipla equilibria, as obtained by Owen et al. (2009). To summarize, accounting for heterogeneity, we can conclude for the existence of two difference of groups in the growth process: one in which countries converge and one in which the positive and significant coefficient associated to the initial level of GDP per capita suggests the lack of convergence and the possible existence of multipla equilibria.

The volatility of growth rate is mainly due to the unemployment rate and to the financial development. This implies that changing in the labor market and in the financial sector are the main causes of the economics, respectively, instability and stability. The high level of financial development is found to be negatively related to the growth variability. This could be due to the direct connection between the financial development and the household consumption. As Aghion et al. (1999), and Easterly et al. (2001) suggest, an increase in the private credit to

GDP generates more consumption smoothness, by reducing the household liquidity constraints; in turn, the less consumption volatility (smoothed by the less liquidity constraints) leads to less volatility in growth. Unemployment is found here to play a destabilizing role on output fluctuation. This could be due to the fact that an increase in the unemployment level generates a decrease in consumption. Inflation, openness to trade and government consumption are found to be non significantly different from zero in the bivariate equation for the scale parameters (see Table 7).

An high level of openness to trade is associated to an improvement in the financial and commercial risk sharing with foreign countries (Cecchetti et al., 2006) and to a consequent increase in the vulnerability to the demand and supply shock (Newbery and Stiglitz, 1984). On the other hand, stabilizing effect of the openness to trade could be due to the financial structure of country itself, i.e. the most exposed to capital flows, the most stabilizing effect on growth openness to trade (Cavallo et al., 2008), or to the degree of diversification of exports (Haddad et al., 2013). Furthermore, we obtain that cyclical fluctuations in the growth rate are negatively related to the labour market participation (Okun, 1962) and to the inflation rate.

5.4 Clustering

An interesting by-product of our approach is the possibility to cluster countries on the basis of their posterior probabilities $w_{ik_1k_2}$. The i-th country can be classified in the $k_1 - k_2$-th group if $\hat{w}_{ik_1k_2} = max_{k_1k_2}(\hat{w}_{i11}, \ldots, \hat{w}_{iK_1K_2})$. It is worth nothing that each group is characterized by homogeneous values of (estimated) random effects; thus, conditionally on observed covariates, countries clustered in the same group share a similar behaviour with respect to the event of interest (i.e. GDP level and growth). This represents a substantial difference with conclusion derived by assuming any parametric approach for the random terms.
Table 8 displays the a posteriori classification. With respect to the GDP level groups, $k_1 = 1$ and $k_1 = 6$ cluster well-developed countries (with any few exceptions), while the poorest countries are clustered in $k_1 = 4$. It is interesting to notice that high levels of GDP are often associate to higher propensity to grow. Indeed, all countries (but Costa Rica, Mexico, Panama, Turkey and Venezuela) clustered in $k_1 = 1$ or $k_1 = 6$ are assigned to $k_2 = 1$, i.e. the growth group with the highest propensity to growth, somehow alleviated by the initial GDP level. Similarly, the “poorest countries” share a lower propensity of economic growth with the exception of China and Thailand (as expected).
The obtained classification is, in this case, not only a mathematical tool able to capture the unobserved heterogeneity, but groups may have a “physical” meaning. Indeed, countries in the same cluster often share similar technological, institutional and/or geographical characteristics (e.g. OECD countries are clustered together), and in general a similar socio-economic background.
A final remark concerns the impact of initial GDP level on growth because it it important to check for convergence. Our results suggest two different process. The first one involves developed countries, whose growth is relatively high and in which higher values of GDP contributes to the growth process, thus leading to “convergence”. On the other hand, for “poorest” countries differences will increase as the initial GDP positively affects economic growth leading to divergence.

6 Conclusion

In this paper we introduce a flexible multivariate multidimensional random model allowing for all model parameters to depend on covariates in a regression framework. We relax the common unidimensionality assumption of the random effects distribution, allowing for a general and flexible association structure among the outcomes. The proposed approach is motivated by the analysis of economic growth in presence of heterogeneous behaviour. We jointly model GDP level and growth by further including a regression model for the variance of growth, to check for the effects of financial variables on the volatility of the growth process. Our empirical findings provide evidence of heterogeneous behaviours in both GDP level and growth rate, confirming the need of a flexible approach to properly reflect all data features. Such heterogeneous behaviours could be due to differences in institutional and technological factors and may contribute to reach (or not) economic convergence. At last, we would remark that estimated covariates effects are in line with the augmented Solow model theory, additionally the growth rate volatility is mainly related to unemployment and financial development. Of course, the model can be extended in several ways. Here, we account for heavy tails in the growth rate distribution, but other distributions than the t one can be considered, as well as approaches to deal with outliers (if any). More than two outcomes can be jointly modelled of the price of a high computational burden involved in the estimation step. An interesting extension would deal with time-varying heterogeneity. Indeed, a limitation of our proposal is that we assume time-constant random effects.

References

[1] Aghion, P., E. Caroli, and C. Garcia-Penalosa (1999): "Inequality and economic growth: the perspective of the new growth theories", *Journal of Economic literature*, pp. 1615–1660.

[2] Aitkin, M. (1999): "A general maximum likelihood analysis of variance components in generalized linear models", *Biometrics*, Vol. 55, No. 1, pp. 117–128.

[3] Alfó, M., G. Trovato, and R. J. Waldmann (2008): "Testing for country heterogeneity in growth models using a finite mixture approach", *Journal of applied Econometrics*, Vol. 23, No. 4, pp. 487–514.

[4] Ardiç, O. P. (2006): "The gap between the rich and the poor: Patterns of heterogeneity in the cross-country data", *Economic Modelling*, Vol. 23, No. 3, pp. 538–555.

[5] Barro, R. J. (1991): "Economic growth in a cross section of countries", *The Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 407–443.

[6] Baştürk, N., R. Paap, and D. van Dijk (2012): "Structural differences in economic growth: an endogenous clustering approach", *Applied Economics*, Vol. 44, No. 1, pp. 119–134.

[7] Bernanke, B. S. and R. S. Gürkaynak (2002): "Is growth exogenous? taking mankiw, romer, and weil seriously", in *NBER Macroeconomics Annual 2001, Volume 16*: MIT Press, pp. 11–72.

[8] Bertarelli, S. and R. Bernardini Papalia (2013): "Nonlinearities in economic growth and club convergence", *Empirical Economics*, Vol. 44, No. 3.

[9] Böhning, D. (1995): "A review of reliable maximum likelihood algorithms for semiparametric mixture models", *Journal of Statistical Planning and Inference*, Vol. 47, No. 1, pp. 5–28.

[10] Bond, S. R., A. Hoeffler, and J. Temple (2001): "Gmm estimation of empirical growth models.".

[11] Caselli, F., G. Esquivel, and F. Lefort (1996): "Reopening the convergence debate: a new look at cross-country growth empirics", *Journal of Economic Growth*, Vol. 1, No. 3, pp. 363–389.

[12] Cavallo, E. A., J. De Gregorio, and N. V. Loayza (2008): "Output volatility and openness to trade: A reassessment [with comments]", *Economia*, pp. 105–152.

[13] Cecchetti, S. G., A. Flores-Lagunes, and S. Krause (2006): "Assessing the sources of changes in the volatility of real growth",Technical report, National Bureau of Economic Research.

[14] Davidson, R. and J. G. MacKinnon (1993): "Estimation and inference in econometrics", *OUP Catalogue*.

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

[16] Doppelhofer, G., R. I. Miller, and X. Sala-i Martin (2000): "Determinants of long-term growth: A bayesian averaging of classical estimates (bace) approach",Technical report, National bureau of economic research.

[17] Durlauf, S. N. and P. A. Johnson (1995): "Multiple regimes and cross-country growth behaviour", *Journal of Applied Econometrics*, Vol. 10, No. 4, pp. 365–384.

[18] Durlauf, S. N. and D. T. Quah (1999): "The new empirics of economic growth", *Handbook of macroeconomics*, Vol. 1, pp. 235–308.

[19] Easterly, W., R. Islam, and J. E. Stiglitz (2001): "Shaken and stirred: explaining growth volatility", in *Annual World Bank conference on development economics*, Vol. 191, p. 211.

[20] Fernandez, C., E. Ley, and M. F. Steel (2001): "Model uncertainty in cross-country growth regressions", *Journal of applied Econometrics*, Vol. 16, No. 5, pp. 563–576.

[21] Ferreira, J. T. S. and M. F. Steel (2006): "A constructive representation of univariate skewed distributions", *Journal of the American Statistical Association*, Vol. 101, No. 474.

[22] Ferreira, J. and M. F. Steel (2004): "Bayesian multivariate regression analysis with a new class of skewed distributions", *Statistics Research Report*, Vol. 419.

[23] Fitzmaurice, G., M. Davidian, G. Verbeke, and G. Molenberghs (2008): *Longitudinal data analysis*: Boca Raton, FL: Chapman and Hall;.

[24] Fraley, C. and A. E. Raftery (2002): "Model-based clustering, discriminant analysis, and density estimation", *Journal of the American Statistical Association*, Vol. 97, No. 458, pp. 611–631.

[25] Giovanni, J. d. and A. A. Levchenko (2009): "Trade openness and volatility", *The Review of Economics and Statistics*, Vol. 91, No. 3, pp. 558–585.

[26] Goetz, S. J. and D. Hu (1996): "Economic growth and human capital accumulation: simultaneity and expanded convergence tests", *Economics Letters*, Vol. 51, No. 3, pp. 355–362.

[27] Haddad, M., J. J. Lim, C. Pancaro, and C. Saborowski (2013): "Trade openness reduces growth volatility when countries are well diversified", *Canadian Journal of Economics/Revue canadienne d'économique*, Vol. 46, No. 2, pp. 765–790.

[28] Hubert, L. and P. Arabie (1985): "Comparing partitions", *Journal of classification*, Vol. 2, No. 1, pp. 193–218.

[29] Islam, N. (1995): "Growth empirics: a panel data approach", *The Quarterly Journal of Economics*, pp. 1127–1170.

[30] Kerekes, M. (2012): "Growth miracles and failures in a markov switching classification model of growth", *Journal of Development Economics*, Vol. 98, No. 2, pp. 167–177.

[31] Lee, K., M. H. Pesaran, and R. P. Smith (1997): "Growth and convergence in a multi-country empirical stochastic solow model", *Journal of applied Econometrics*, Vol. 12, No. 4, pp. 357–392.

[32] Levine, R. and D. Renelt (1992): "A sensitivity analysis of cross-country growth regressions", *The American economic review*, pp. 942–963.

[33] Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A contribution to the empirics of economic growth",Technical report, National Bureau of Economic Research.

[34] Martínez-Zarzoso, I. and A. Maruotti (2011): "The impact of urbanization on co¡ sub¿ 2¡/sub¿ emissions: Evidence from developing countries", *Ecological Economics*, Vol. 70, No. 7, pp. 1344–1353.

[35] Newbery, D. M. and J. E. Stiglitz (1984): "Pareto inferior trade", *The Review of Economic Studies*, Vol. 51, No. 1, pp. 1–12.

[36] Okun, A. M. (1962): "The predictive value of surveys of business intentions", *The American Economic Review*, pp. 218–225.

[37] Owen, A. L., J. Videras, and L. Davis (2009): "Do all countries follow the same growth process?", *Journal of Economic Growth*, Vol. 14, No. 4, pp. 265–286.

[38] Pesaran, M. H. and R. Smith (1995): "Estimating long-run relationships from dynamic heterogeneous panels", *Journal of econometrics*, Vol. 68, No. 1, pp. 79–113.

[39] Pittau, M. G., R. Zelli, and P. A. Johnson (2010): "Mixture models, convergence clubs, and polarization", *Review of Income and Wealth*, Vol. 56, No. 1, pp. 102–122.

[40] Solow, R. M. (1956): "A contribution to the theory of economic growth", *The quarterly journal of economics*, pp. 65–94.

[41] Temple, J. (1999): "The new growth evidence", *Journal of economic Literature*, pp. 112–156.

[42] Temple, J. (2000): "Growth regressions and what the textbooks dont tell you", *Bulletin of Economic research*, Vol. 52, No. 3, pp. 181–205.

[43] Verbeke, G., S. Fieuws, G. Molenberghs, and M. Davidian (2014): "The analysis of multivariate longitudinal data: A review", *Statistical methods in medical research*, Vol. 23, No. 1, pp. 42–59.

[44] Wang, Y. (2010): "Maximum likelihood computation for fitting semiparametric mixture models", *Statistics and Computing*, Vol. 20, No. 1, pp. 75–86.

[45] Winkelmann, R. (2000): "Seemingly unrelated negative binomial regression", *Oxford Bulletin of Economics and Statistics*, Vol. 62, No. 4, pp. 553–560.

[46] Wooldridge, J. (2002): "Econometric analysis of cross section and panel data. Massachusetts".

[47] Zellner, A. (1962): "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias", *Journal of the American statistical Association*, Vol. 57, No. 298, pp. 348–368.

Table 1: Simulation results: Scenario 1.

	True	Estimate	Bias	Std. dev.
		n=100, T=5		
$u_{k_1=1}$	-1.00	-1.020	-0.020	0.265
$u_{k_1=2}$	1.00	1.012	0.012	0.265
λ_{11}	0.50	0.505	0.005	0.111
$u_{k_2=1}$	2.00	2.012	0.012	0.306
$u_{k_2=2}$	-2.00	-2.005	-0.005	0.243
λ_{12}	0.50	0.494	-0.006	0.149
γ_{01}	0.50	0.489	-0.011	0.072
γ_{11}	0.75	0.760	0.010	0.120
γ_{02}	1.00	0.991	-0.009	0.068
γ_{12}	0.25	0.253	0.003	0.122
π_{11}	0.40	0.420	0.020	0.048
π_{12}	0.10	0.090	-0.010	0.049
π_{21}	0.20	0.196	-0.004	0.047
π_{22}	0.30	0.294	-0.006	0.063
	Average Rand Index= 0.800			
		n=100, T=10		
$u_{k_1=1}$	-1.00	-1.006	-0.006	0.139
$u_{k_1=2}$	1.00	1.000	0.000	0.142
λ_{11}	0.50	0.500	0.000	0.078
$u_{k_2=1}$	2.00	2.017	0.017	0.171
$u_{k_2=2}$	-2.00	-2.004	-0.004	0.134
λ_{12}	0.50	0.495	-0.005	0.098
γ_{01}	0.50	0.502	0.002	0.049
γ_{11}	0.75	0.741	-0.009	0.086
γ_{02}	1.00	0.993	-0.007	0.046
γ_{12}	0.25	0.257	0.007	0.078
π_{11}	0.40	0.407	0.007	0.052
π_{12}	0.10	0.096	-0.004	0.034
π_{21}	0.20	0.196	-0.004	0.045
π_{22}	0.30	0.300	0.000	0.041
	Average Rand Index= 0.905			

Table 2: Simulation study: Scenario 2

	True	**Estimate**	**Bias**	**Std. dev.**	**Estimate**	**Bias**	**Std. dev.**
		n=100, T=5			**n=100, T=10**		
$u_{k_1=1}$	-1.00	-1.028	-0.028	0.337	-1.007	-0.007	0.160
$u_{k_1=2}$	1.00	1.035	0.035	0.252	1.016	0.016	0.123
λ_{11}	0.50	0.498	-0.002	0.111	0.499	-0.001	0.074
$u_{k_1=1}$	2.00	2.200	0.200	0.715	2.071	0.071	0.403
$u_{k_2=2}$	-2.00	-2.271	-0.271	0.935	-2.090	-0.090	0.401
$u_{k_2=3}$	0.00	-0.136	-0.136	0.746	-0.066	-0.066	0.616
λ_{12}	0.50	0.498	-0.002	0.150	0.504	0.004	0.097
γ_{01}	0.50	0.490	-0.010	0.071	0.496	-0.004	0.049
γ_{11}	0.75	0.755	0.005	0.123	0.751	0.001	0.084
γ_{02}	1.00	0.989	-0.011	0.079	0.993	-0.007	0.050
γ_{12}	0.25	0.252	0.002	0.014	0.255	0.005	0.084
π_{11}	0.10	0.038	-0.062	0.048	0.038	-0.062	0.048
π_{12}	0.10	0.129	0.029	0.061	0.129	0.029	0.061
π_{13}	0.20	0.230	0.030	0.058	0.230	0.030	0.058
π_{21}	0.20	0.191	-0.009	0.075	0.191	-0.009	0.075
π_{22}	0.30	0.360	0.060	0.070	0.360	0.060	0.070
π_{23}	0.10	0.051	-0.049	0.058	0.051	-0.049	0.058
		Average Rand Index= 0.740			Average Rand Index= 0.841		
	True	**Estimate**	**Bias**	**Std. dev.**	**Estimate**	**Bias**	**Std. dev.**
		n=1000, T=5			**n=1000, T=10**		
$u_{k_1=1}$	-1.00	-0.999	0.001	0.097	-1.000	0.000	0.050
$u_{k_1=2}$	1.00	1.000	0.000	0.073	1.002	0.002	0.039
λ_{11}	0.50	0.501	0.001	0.033	0.501	0.001	0.025
$u_{k_2=1}$	2.00	2.054	0.054	0.215	2.005	0.005	0.100
$u_{k_2=2}$	-2.00	-2.039	-0.039	0.358	-2.005	-0.005	0.084
$u_{k_2=3}$	0.00	-0.016	-0.016	0.542	-0.002	-0.002	0.185
λ_{12}	0.50	0.500	0.000	0.047	0.499	-0.001	0.031
γ_{01}	0.50	0.500	0.000	0.022	0.500	0.000	0.015
γ_{11}	0.75	0.749	-0.001	0.037	0.751	0.001	0.026
γ_{02}	1.00	1.000	0.000	0.021	0.999	-0.001	0.015
γ_{12}	0.25	0.249	-0.001	0.038	0.250	0.000	0.026
π_{11}	0.10	0.072	-0.028	0.038	0.086	-0.014	0.014
π_{12}	0.10	0.129	0.029	0.032	0.111	0.011	0.015
π_{13}	0.20	0.213	0.013	0.032	0.203	0.003	0.021
π_{21}	0.20	0.198	-0.002	0.041	0.199	-0.001	0.021
π_{22}	0.30	0.299	-0.001	0.054	0.299	-0.001	0.023
π_{23}	0.10	0.090	-0.010	0.040	0.101	0.001	0.026
		Average Rand Index= 0.774			Average Rand Index= 0.859		

Table 3: Summary statistics

	Mean	**Std. Dev.**	**Variable Description**	**Sources**
GDP level				
sk	0.002	0.001	share of output invested in physical capital	PWT 8.0
sh	0.632	0.34	share of output invested in human capital	World Bank
$ng\delta$	0.067	0.012	population growth rate + 0.05$^{(*)}$	PWT 8.0
lnyc	8.509	1.268	log of income per capita	PWT 8.0
Growth				
unemp	0.612	0.077	unemployment rate	PWT 8.0
infl	0.519	0.312	log of consumer price	PWT 8.0
open	66.7	38.05	openness to trade	World Bank
govcons	15.329	5.853	government consumption (as share of GDP)	World Bank
fin	45.656	39.801	domestic credit on GDP	World Bank
N		519		

Notes: (*): 0.05 is the commonly used value for approximating the depreciation growth rate and the technological rate.

Table 4: Respone Variables: Summary statistics

	Mean	Std. Dev.	Skewness	Kurtosis	Min	Max	N
GDP level	8.6	1.3	-0.27	1.96	5.42	10.70	519
GDP growth	0.9	0.2	-0.37	10.47	-1.33	1.31	519

Table 5: Penalized Likelihood Criteria: Univariate model

	LLK	AIC	BIC
$K_1 = 2$	-360.89	735.77	754.08
$K_1 = 3$	-312.95	643.89	667.43
$K_1 = 4$	-277.54	577.08	605.85
$K_1 = 5$	-265.11	556.22	590.22
$K_1 = 6$	-248.04	**526.08**	**565.31**
$K_1 = 7$	-258.81	551.62	596.08
	LLK	AIC	BIC
$K_2 = 2$	172.22	**-322.44**	**-293.67**
$K_2 = 3$	172.24	-316.47	-279.86
$K_2 = 4$	173.19	-312.37	-267.91
$K_2 = 5$	173.18	-306.36	-254.06

Table 6: Penalized Likelihood Criteria: Bivariate model

K_1	K_2	llk	AIC	BIC
2	2	-187.32	414.64	466.94
2	3	-186.55	421.1	483.86
2	4	-185.51	427.02	500.24
2	5	-185.57	435.14	518.82
2	6	-184.91	441.82	535.96
3	2	-147.21	340.42	400.57
3	3	-136.14	328.28	401.50
3	4	-152.45	370.9	457.20
3	5	-141.97	359.94	459.31
3	6	-134.52	355.04	467.49
4	2	-97.18	246.36	314.35
4	3	-96.29	256.58	340.26
4	4	-93.43	262.86	362.23
4	5	-91.77	271.54	386.61
4	6	-85.44	270.88	401.64
5	2	-66.97	191.94	267.78
5	3	-57.64	187.28	281.42
5	4	-55.25	196.5	308.95
5	5	-54.27	208.54	339.30
5	6	-66.52	247.04	396.10
6	2	-50.42	164.84	**248.52**
6	3	-39.72	**159.44**	264.04
6	4	-36.47	168.94	294.47
6	5	-61.48	234.96	381.41
6	6	-35.29	198.58	365.95

Table 7: Results

	Univariate		Bivariate	
	Coef.	SE	Coef	SE
Income per capita:				
μ_{it1}				
sk	0.07 **	0.03	0.14 ***	0.03
sh	0.72 ***	0.02	0.46 ***	0.03
$ng\delta$	-0.33 ***	0.10	-0.61 ***	0.1
$u_{0k_1=1}$	8.46 ***	0.32	9.64 ***	0.31
$u_{0k_1=2}$	9.02 ***	0.06	7.48 ***	0.31
$u_{0k_1=3}$	9.59 ***	0.06	8.07 ***	0.3
$u_{0k_1=4}$	10.19 ***	0.06	6.97 ***	0.31
$u_{0k_1=5}$	10.84 ***	0.06	8.59 ***	0.3
$u_{0k_1=6}$	11.29 ***	0.09	9.01 ***	0.3
$log(\sigma_{it1})$				
γ_{01}	-1.23 ***	0.03	-1.28 ***	0.03
Observations	519		519	
K_1	6		6	
$\ell^{(*)}$	-248.81			
ℓ			-50.42	
Growth rate:				
μ_{it2}				
$u_{0k_2=1}$	-0.01	0.05	1.05 ***	0.12
$u_{0k_2=2}$	1.11 ***	0.15	-0.1	0.09
$lnyc_{k_2=1}$	0.01 **	0.01	-0.09 ***	0.12
$lnyc_{k_2=2}$	-0.1 ***	0.02	0.02 **	0.01
$log(\sigma_{it2})$				
γ_{02}	-1.53 ***	0.52	-1.52 ***	0.51
$unemp$	1.17 **	0.55	1.34 **	0.53
$infl$	0.01	0.08	0.03	0.08
$open$	0.05	0.08	0.05	0.08
$govcons$	-0.11	0.11	-0.15	0.11
fin	-0.31 ***	0.05	-0.31 ***	0.05
ν_{it2}				
$\tilde{\gamma}$	1.72 ***	0.17	1.69 ***	2.23
Observations	519		519	
k_2	2		2	
$\ell^{(*)}$	172.22			
ℓ			-50.42	

Significance level: $***: 0.1\%$ $**: 1\%$ $*: 5\%$
Notes: $\ell^{(*)}$: log-likelihood for the univariate model, ℓ: log-likelihood for the bivariate model. Dependent variables: 5 years forward value of log of GDP per capita (top of the Table), and 5 years forward value of growth rate.

Table 8: Clustering results

	K_2	
K_1	1	2
1	Australia, Austria, Belgium, Canada, Czech Rep., Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, Trinidad & Tobago, UK, USA	
2		Bangladesh, Benin, Burkina Faso, Burundi, Rep. Congo, India, Kenya, Madagascar, Mali, Moldova, Niger, Rwanda, Sri Lanka, Syria, Tanzania, Uganda
3	China	Bolivia, Cameroon, Chad, Djibouti, Egypt, Honduras, Indonesia, Jamaica, Jordan, Mauritania, Morocco, Pakistan, Paraguay, Peru, Phillippines, Senegal, Sierra Leone
4		Rep. Central African, Rep. Dem. Congo, El Salvador, Malawi, Mozambique, Nepal, Nigeria, Togo
5	Thailand	Bulgaria, Colombia, Dominican Rep., Ecuador, Guatemala, Serbia, South Africa, Tunisia, Uruguay,Zimbabwe
6	Angola, Argentina, Botswana, Chile, Croatia, Estonia, Greece, Hungary, Rep. of Korea, Latvia, Malaysia, Maldives, Mauritius, Poland, Poland, Portugal, Romania, Russia, Slovakia	Costa Rica, Mexico, Panama, Turkey, Venezuela

Figure 1: Histograms of response varialbes

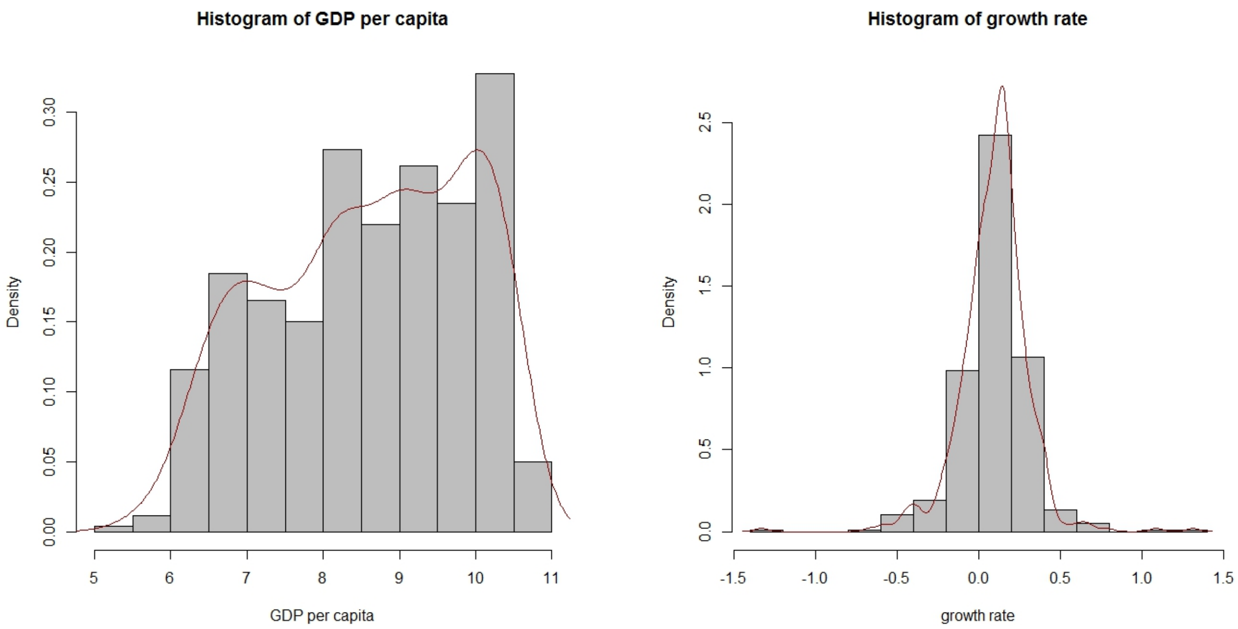
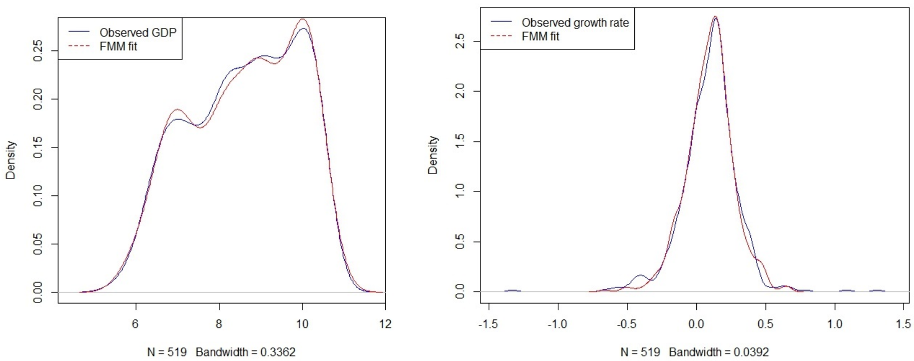


Figure 2: Model fitting: GDP level (left box), GDP growth (right box)



Chapter 2

Estimating the determinants of growth stability and instability in Sub - Saharan African countries: a Markov Switching Approach

Alessandra Marcelletti*,† Igbinoba Emmanuel,‡ Giovanni Trovato§

Abstract

The goal of the present paper is to study the asymmetries characterizing African growth path. To model switching between output collapses and growth stability, we apply a Markov-switching model with time varying transition probabilities on 37 African countries over the period 1987-2011. Two distinct growth regimes are identified: a stable growth regime, which, despite the presence of negative values of the growth rate, is stable, and a highly volatile growth regime, in which the growth behaviour range from negative to positive picks of the growth rates, and vice versa. From our analysis, we observed that the likelihood of an economy being in a stable growth regime increases with the political regime, trade and variation in the exchange rate.

JEL classification: C24, F41, F43.

Keywords: Markov Regime-Switching Model; Economic Growth; Sub-Saharan Africa.

1 Introduction

The present paper aims to study the change of growth patterns and its determinants within African countries. To do that, we apply a Markov Switching Model with time varying transition probability, relying on the idea that the country's growth path is the result of different growth regimes (Jerzmanowski, 2006).
In other words, we assume a latent trait underlying the time series of growth rate, which explains the transition between different growth regimes occurring without obvious changes in country-fundamentals.
The lack of a linear pattern is typical of developing and emerging countries, that are characterized by large fluctuations and swings of the growth rate (see among others Becker et al., 2006), and asymmetric interchange between phases of growth acceleration and regime collapse (e.g. Easterly et al., 1993; Jerzmanowski, 2006; Kerekes, 2012; Jones and Olken, 2008). In particular, the most volatile growth rate and the most frequent output collapse[1] are recorded in

*Corresponding Author. Dipartimento di Economia Diritto ed Istituzioni, Università di Roma "Tor Vergata" via Columbia 2 00133 (RM) Italy e-mail: marcelletti@economia.uniroma2.it

†Dipartimento di Scienze Politiche, Luiss Guido Carli, Viale Romania 32, 00197, Roma, email: amarcelletti@luiss.it"

‡Dipartimento di Economia Diritto ed Istituzioni, Università di Roma "Tor Vergata" via Columbia 2 00133 (RM) Italy

§Dipartimento di Economia e Finanza, Università di Roma "Tor Vergata" via Columbia 2 00133 (RM) Italy

[1] Output collapse is described as a dip or a decline in GDP.

Sub-Saharan Africa. This region is characterized by different growth behaviours over time, in terms of different balanced growth paths, different within-state variability and different long-run growth rates. This stylized fact implies a failure, for this region, to conform to economic growth theories, especially of convergence model, that predicts a rapid and regular growth rate for these countries. Indeed, the difficulty for African countries to sustain growth for substantial period, and asymmetric switching between stable growth and economic collapses is tested by Byrne (2010)[2]. She also provides empirical evidence of different behaviours of the time series of growth over phases of stable growth and output collapse.
To model the evolution of the growth rate over time accounting for the different regimes, we apply a Markov Switching approach with time varying transition probability (e.g. Filardo, 1994, Filardo and Gordon, 1998, Diebold et al., 1994 Kerekes, 2012) to 37 African countries over the period 1987–2011. This statistical approach allows us to disentangle the different growth paths within countries, and to reject the hypothesis, commonly followed by standard growth regression, that countries follow constantly and indefinitely a linear growth path (Pritchett, 2000).
Entering into details, it allows us to deal with non-linear properties of the regression function such as asymmetry, dependency weights and volatility. Furthermore, it splits the time series into a discrete number of regimes, simultaneously estimating the transition from one state to another, whether the variable leading to regime shift is unobservable. Although other models commonly used to study the growth rate variation, such as autoregressive (AR), moving average (MA) or autoregressive integrated moving average (ARIMA) can capture the dynamics in data, they are unable to capture the fact that the variable under estimation change its behaviour during time, i.e. it switches between regimes (see among others Kuan, 2002, Bilgili et al., 2012).
The time varying transition probability allows us to deal with two challenges. Firstly, we can understand long run growth pattern by determining the length of time a country remains in the same regime as dependent on country fundamentals. For this reason, we believe that the Hamilton (1989)'s seminal paper is too restrictive for explaining economic growth dynamics, and we apply the Filardo (1994) and Diebold et al. (1994) extension. Secondly, since the transition probability depends on country-specific factors, we allow for cross country differences in the growth process, avoiding to model the growth process as common across countries. In this way, we can evaluate the effect of country fundamentals on switching regimes, i.e. the impact of economic factors on the movement across growth phases.
Previous works about switching growth determinants are also present in literature with special references to papers by Hausmann et al. (2005), Byrne (2010) and Jerzmanowski (2006). The present paper closely follows the Byrne (2010)'s intuition about the need for applying Markov Switching Model with time varying transition probabilities to understand the African countries growth dynamics. On the other hand, it takes its theoretical basis from the Hausmann et al. (2005)'s work for the choice of the explanatory variables of the discrete latent process. Their empirical findings provide evidence in favour of the accelerating mechanisms of investment, trade, and real exchange rate depreciation and changes to political regimes[3]. It is worth noting that empirical evidence shows that expansion of international trade has been noted to be a primary catalyst for starting growth phases (Becker et al., 2006).

[2]Byrne (2010) applies a Markov Switching Model with time varying transition probability to a panel of Sub-Saharan African countries over the period 1960–2004. She models the transition probability matrix as dependent on education, quality of institution, trade, and dummy variables capturing the composition of output (fuel, agriculture, manufacture).

[3]Hausmann et al. (2005) identify the growth acceleration phases defining a threshold for the growth rate.

From our analysis, two distinct growth regimes are identified: a stable growth regime, which despite the presence of negative values of the growth rate, is stable and a highly volatile growth regime, in which the growth behaviour range from negative to positive picks of the growth rates, and the variance within the group is extremely high. Furthermore, the probability of switching to the highly volatile regime positively depends on exchange rate depreciation, trade and foreign direct investment.
The paper is structured as follows. Section 2 describes the Markov Switching approach with time varying transition probabilities. Section 3 presents the empirical application on African countries, pointing out the results. Section 4 concludes and presents some possible extensions.

2 Statistical Specification: Two-State Markov switching approach with explanatory variables

To formalize the different phases of growth each country faces over time, we start defining the existence of M latent states of the economy. In particular, since African countries are mainly characterized by two asymmetric regimes, unstable and stable growth (Byrne, 2010), we specify the two-state Markov switching approach. Furthermore, to allow the growth process to be not identical across countries, we model the switching process as dependent on country specific variables, through the time varying transition probability.
Formally, we denote S_t the discrete latent variable that takes only two values over the space M, such that $S_t = \{1, 2\}$, and $\mathbf{y}_t = \{y_{it}\}$ the vector containing continuous realization of the dependent variable (i.e. the growth rate) recorded for each country $i \quad (i = 1, \ldots, n)$ at time $t \quad (t = 1, \ldots, T)$. We assume that the dependent variable follows an AR(1) process as:

$$y_{it} = \alpha_{S_t} + \beta_{S_t} y_{it-1} + \varepsilon_{S_t it} \tag{1}$$

where $\varepsilon_{S_t it} \sim N(0, \sigma^2_{S_t})$, with $cov(\varepsilon_{1,it}, \varepsilon_{2,it}) = 0$, is the error term which allows for different variability within the states, i.e. the regime varying variance represents the uncertainty measure of output occurring in each state of the economy. The evolution of the dependent variable over time depends on the history of the variable itself y_{it-1}. The autocorrelation effect, captured by the $\boldsymbol{\beta}_{S_t}$ parameter, changes its magnitude on the basis of which latent regime occurs, and it allows for formalizing the impact of the latent variable on the growth path.
Moreover, we explicit the two different paths of the growth rate according to which state of the economy is verified as:

$$\begin{cases} y_{it} = \alpha_1 + \beta_1 y_{it-1} + \varepsilon_{it1} \quad if \quad S_t = 1 \\ \\ y_{it} = \alpha_2 + \beta_2 y_{it-1} + \varepsilon_{it2} \quad if \quad S_t = 2 \end{cases} \tag{2}$$

The model described by the system of equation (2) highlights the key feature of the Markov Switching approach: the definition of the conditional distribution of the time series of the response variable as dependent on the underlying latent state[4]. We specify also the conditional

[4]It is worth noting that the formalized model above may also be extended by including a set of $\boldsymbol{X}_t$ explanatory variables (see Kim et al., 2008 for further details.) affecting the distribution of the growth rate. This paper aims to understand what leads the transition between different growth regime. Thus, for sake of simplicity and parsimony in the computational framework, we reduce the model on the "direct" effect on growth dynamics only as dependent on the previous value of the growth rate.

probability density of the time series vector $\mathbf{y}_t$, given the two state-switching model as:

$$P(\mathbf{y}_t|\mathbf{y}_{t-1},\mathbf{z_t})\begin{cases} f(\mathbf{y}_t|\mathbf{y}_{t-1},\boldsymbol{\theta}_1) & if \quad S_t=1 \\ \\ f(\mathbf{y}_t|\mathbf{y}_{t-1},\boldsymbol{\theta}_2) & if \quad S_t=2 \end{cases} \tag{3}$$

Following Diebold et al. (1994) and Filardo (1994), we assume that the latent variable is not completely unobservable, but it partially depends on economic factors included in the set of explanatory variables $\mathbf{z_t}$. To endogenize probabilistic changes of regime, since the statistical properties of the discrete latent variable are summarized by the transition matrix, we define the time varying transition probability $\mathbf{\Pi}(\mathbf{z}_t) = p(S_t = j|S_{t-1} = k, \mathbf{z}_t) = p_{jk}(\mathbf{z_t})$, with $\{j, k\} \in M$, as dependent on economic and political factors. We denote $p_{jk}(\mathbf{z_t})$ the probability of having in place regime j in period t, conditional on having in place regime k in period $t-1$. By specifying that the switching of regimes follows a first order Markov chain, as $p_{11} = p(S_t = 1|S_{t-1} = 1, \mathbf{z}_t) = p(\mathbf{z}_t)$ and $p_{22} = p(S_t = 2|S_{t-1} = 2, \mathbf{z}_t) = q(\mathbf{z}_t)$, we infer the influence of $\mathbf{z}_t$ on the transition matrix $\mathbf{\Pi}(\mathbf{z_t})$, by a probit specification for S_t over $\{1, 2\}$. Thus, we specify the linear regression of the latent variable on the set of covariates and latent traits as:

$$S_t = a_{S_{t-1}} + \mathbf{z}_t'\mathbf{b}_{S_{t-1}} + u_t \tag{4}$$

where $u_t \sim N(0, 1)$, and it is not correlated with ε_{it} (see Filardo and Gordon, 1998). We emphasize that the probability of switching over the state, as well as remaining in the same state, partially depends on time and on variables contained in $\mathbf{z_t}$. Hence, we incorporate economic variables as determinants of the probability of changing or moving from one state of the economy to another (Diebold et al., 1994). The estimation of the full parameter vectors $\boldsymbol{\theta}_{S_t} = (\boldsymbol{\beta}_{S_t}, \sigma^2_{S_t}, \mathbf{\Pi}(\mathbf{z_t}))$, is obtained by applying the Maximum Likelihood Method to the complete data log-likelihood[5]. According to model assumptions, the complete data likelihood function is given by:

$$L(\cdot) = \prod_{t=1}^{T} f(\mathbf{y_t}|\mathbf{y_{t-1}}, \mathbf{z_t}; \boldsymbol{\theta}_{\mathbf{S_t}}) \tag{5}$$

where, for sake of simplicity, we denote $f(\mathbf{y}_{jt}|\mathbf{y}_{t-1}, \mathbf{z}_t; \boldsymbol{\theta}_{S_t})$ as:

$$f(\mathbf{y}_t|\mathbf{y}_{t-1}, \mathbf{z}_t; \boldsymbol{\theta}_{S_t}) = \sum_{j=1}^{M}\sum_{k=1}^{M} f(\mathbf{y}_{jt}|S_t = j, S_{t-1} = k, \mathbf{y}_{t-1}, \mathbf{z}_t; \boldsymbol{\theta}_{S_t})P(S_t = j, S_{t-1} = k|\mathbf{y}_{t-1}, \mathbf{z}_t; \boldsymbol{\theta}_{S_t}) \tag{6}$$

We justify the use of time varying in place of time fixed transition probability by the increased and the ability of the former in capturing systematic changes before and after the turning points. Furthermore, since we include economic variables in the estimation of the switching process, it is possible to evaluate how the end of a highly unstable growth regime is affected by the country fundamentals.

3 Empirical formulation

3.1 Data

The sample is composed by an unbalanced panel of 37 African countries, from 1987 to 2011. Data are retrieved from Summer and Heston database. The dependent variable is the growth

[5] The model is estimated through the Matlab Package developed by Ding (2012)

rate of income per capita. Skewness and kurtosis values (see Table 1), as well as the Shapiro and Wilk test ($W = 0.97290$, with $p = 0.0000$) performed on the growth rate, display departure from normality distribution. As the QQ-Plot confirms (see Figure 1), it is possible to conclude that the growth rate has not a Gaussian distribution. Indeed, the heavy tails suggests that is better to approximate the distribution of the dependent variable through a t-student.
The independent variable is the one year lagged value of the growth rate (see among others Kerekes, 2012; Byrne, 2010).
The variables affecting the first order Markov chain are chosen on the basis of Hausmann et al., 2005 empirical findings, and they are: trade, foreign direct investment, variation in the exchange rate and political regime. Statistical details about transition probability's explanatory variables are provided in Table 2.
It is worth emphasizing that, for sake of simplicity in the coefficient interpretation, the political regime is formalized as a categorical variable, equal to 1 in democracy, 2 in autocracy and 3 in the Transitional Regime. The latter includes: rebel regime, civil war and transitional regime; while different "levels" of autocracy are included in the second category[6].

3.2 Theoretical Specification

In order to understand the variation in the growth path of African countries, following Jerzmanowski (2006), we assume that economic growth is the result of switching between distinct growth regimes. Variables used to model the transition between different states are based on Hausmann et al. (2005)'s empirical findings. Their results provide evidence that foreign direct investment (FDI henceforth), trade, depreciation of the real exchange rate and changes in political regimes are the main reasons behind the phases of economic acceleration. Thus, the regression function can be formalized as[7]:

$$E(gc_{it}) = \alpha_{S_t} + \beta_{S_t} growth_{it-1} \tag{7}$$

where

$$S_t = a_{t-1} + b_{1,S_{t-1}} trade + b_{2,S_{t-1}} political + b_{3,S_{t-1}} fdi + b_{4,S_{t-1}} \Delta exch\_rate \tag{8}$$

Equation (7) and equation (8) model the cross-country differences in growth dynamics, i.e. the non-linear country - specific growth paths, as dependent on political and economic factors[8]. In other words, these equations disentangle the time series of the growth rate according to which latent state occurs in the economy, that, in turn, partially depends on FDI (fdi), trade ($trade$), exchange rate variation ($\Delta exch\_rate$), and political regime ($political$).
Entering into details, we assume that switching in growth dynamics is associated to variables that capture and affect the ability of each country to deal with internal or external shocks. For this reason, political regime is the key variable in our analysis. Indeed, it is empirically demonstrated that the presence of internal conflicts reduces the ability of a country to deal with the impact of shocks (Rodrik, 1999), leading the growth rate to switch between different phases. It is worth noting that as Durlauf et al. (2005) point out, it is common to assume that whereas output collapses are common and significant in terms of magnitude, they should be

[6] The second category of Political Regime includes: One-Party; military state; one party military state; monarchical state; limited multiparty

[7] As specified in Section 2, we assume that within each regime economic growth evolves according to an AR(1) (see among others Kerekes, 2012).

[8] We emphasize that the impact of the $\boldsymbol{\beta}$ parameters of equation (8) should not be read in terms of impact of the economic and political variables on the growth rate, but on the transition from one regime to another.

better explained by the institutional quality rather than changes in political setting. However, as Kerekes (2012) argues, the dependence assumption of the transition probability matrix on the country's quality of institutions has some drawbacks. Firstly, the fact that it is usually estimated at the end of the sample period could generate endogeneity problems; secondly, it could change as consequence of economic growth (Glaeser et al., 2004), leading to inconsistency problem of the Markov Switching model (Kim, 2004).

Furthermore, economic instability within African countries is directly related to FDI, trade and exchange rate. Firstly, FDI are found to be one of the main reason behind economic acceleration, since they boost economic growth through different channels: employment creation, technological know-how, managerial skills, and competitiveness (see among others Kobrin, 2005 and de Mello Jr, 1997 for a survey literature). Furthermore, as Asiedu (2002) points out, the low level of FDI inflows recorded in Africa in past decades is mainly due to political and macroeconomic instability, in terms of poor infrastructure, weak investment promotion strategies and high protectionism (Dupasquier and Osakwe, 2006). Secondly, both trade and exchange rate variation capture economic external shocks, and are crucial for growth phases (Cuaresma and Wörz, 2005). In particular, the sensitivity to trade volatility is higher in less developed countries rather than developed economies (Blattman et al., 2004), while the depreciation of exchange rate has a positive impact on growth if it is allowed to operate through aggregate supply channels (Thapa, 2002; Golley and Tyers, 2006 and Akinbobola and Oyetayo, 2010[9]).

3.3 Results

In our benchmark specification, we assume that there is an unobservable exogenous random process which determines the behaviour of African economic growth rate. Thus, we estimated a two state Markov Switching model with time fixed transition probability. Following Hamilton (1989), we assume that the transition probabilities are constant over time, and they do not depend on covariates[10]. Hence, the benchmark model reads as:

$$E(growth_{it}) = \alpha_S + \beta_S growth_{it-1} \tag{9}$$

where the state (latent) variable S, affecting the intercept and the autocorrelation coefficient, is completely unobservable. It is worth reminding that the use of the time varying (TVTP henceforth) in place of the time fixed transition probability (TF henceforth) is justified by the ability of the former in capturing the impact of economic fundamentals on the turning points. Furthermore, penalized likelihood criteria (AIC and BIC) confirm the choice of the TVTP approach in place of the Hamilton (1989) approach (see Table 5). For sake of completeness, a model with $M = 3$ is performed. However, by looking at the BIC and AIC values[11], the model with two states is chosen in place of the model with three regimes.

Table 5 displays results for the benchmark model (TF) and the TVTP model.

In the TF specification, we find substantial differences between the two regimes, meaning that growth rate of African countries does not follow a unique path. Indeed, State 1 is characterized by a high degree of correlation between the growth rate and its value in the previous period,

[9] Akinbobola and Oyetayo (2010) concluded that the positive effect of exchange rate variation can be within the context of a broader program of adjustments and reforms

[10] The transition probability in the Hamilton (1989)'s specification are formally defined as $p_{11} = p(S_t = 1|S_{t-1} = 1) = p$ and $p_{22} = p(S_t = 2|S_{t-1} = 2) = q$.

[11] Log Likelihood and the BIC value for the $M = 3$ are respectively equal to $-1711,1$ and $3695,3$, with 42 parameters to be estimated.

and by a low level of variability within the State. On the other hand, State 2 has a lower level of autocorrelation on the dependent variable and a high level of volatility within the regime. Lastly, the probability of a country in remaining in the same regime is quite high, meaning that both regimes have very high levels of persistence, whilst there is evidence about the different time length of the regimes. In particular, State 2 is found to last longer than State 1 (respectively, 16.83 time periods and 6.14 periods), and the probability of remaining in State 2 is higher than the one of remaining in State 1. Thus, Markov Switching with TF shows that African countries are most likely of following an unstable growth rate, characterized by picks both in negative and positive terms.
Table 5 displays also results for the model as it is formalized in Section 2. Similar conclusions with respect to the benchmark model could be drawn about the autocorrelation coefficients in the two states, whilst the β_{S_t} is lower in the TVTP model than in the TF, particularly in the first regime. Moreover, the within state variability is observed to be almost the double with respect to the TF model.
As it is formalized in Section 2, the model considered here allows the disturbance term $\varepsilon_{it} \sim N(0, \sigma^2(S_t))$ to vary among groups. This implies the presence of a regime varying variance on the disturbance term, which accounts for the different output volatility occurring in the recession phases with respect to the one in expansion phases. As it is clear from Table 5, both regimes are characterized by high levels of variability which strongly suggests that growth rates in SSA tend to be triggered by different conditions. However, regime 2 is characterized by a higher level of variability, and by looking at the distribution of the observed growth rate in Regime 2, it is clear that countries belonging to this regime are more volatile in terms of growth experience. High output volatility could be associated with negative aspects of underdevelopment and empirical evidence points out that high volatility has negative effects on growth or is at least closely associated with lower growth. Causes of these high volatility in SSA can be broadly classified into two groups: those associated with higher exposure to exogenous shocks and augmenting factors, and those related to faulty policies and structural issues. Same conclusion could be drawn for the endogenous variables affecting the transitional probability matrix, with the exception of the exchange rate.
Lastly, the probability of remaining in each regime is very high (almost 96%) meaning both regimes have very high levels of persistence. There is no evidence about the prevalence of one state to another. The posterior analysis of our data confirms that both groups are characterized by the same expected durability. A possible cause for the high persistence levels might be related to country characteristics such as education levels or political stability as well as shocks, particularly those due to terms of trade
As a non-parametric approach, the goodness of fit is checked by graphical tools, as the empirical cumulative density function and the kernel distribution. As both Figure 2 and Figure 3 show, the Data Generating Process is well approximate by the Markov Switching Model with TVTP. It is worth noting that the higher the probability threshold used to compute the predicted value, the better the goodness of fit.

3.3.1 Regime classification

As a by product of the model, two different groups of countries are found. As the empirical findings show, the two groups differ above all because of their variability. Indeed, countries belonging to State 1 are characterized by sustained growth rate, while the second group is characterized by highly volatile growth rates. This is also confirmed by the growth rate distribution. Indeed, it is noted that when country experience sudden changes in the growth rate,

i.e. when there is a pick in the growth rate, the probability of belonging to State 2 is greater than 70% (e.g. by looking for example at the growth rate distribution of Lesotho, Malawi). Nevertheless, countries belonging to both states strongly depend on primary exports, political and economic instability, demographic changes (rapid population growth) and social conditions (ethnolinguistic and religious diversity) which represent serious obstacle to growth. The main determinants of economic growth variability in both states are geographical factors. Indeed, climate, soil, topography and disease ecology makes African countries suffer from chronically low agricultural productivity, high disease burden and very low level of international trade.
As Table 5 displays, the duration of a highly volatile economy transiting to sustained growth is higher than the duration of an economy with sustained growth retaining its growth path (expected duration of the first state is 6.55 time period, against 8.07 in the second state). Thus, countries with highly volatile growth tend to experience longer periods of output collapse and they are likely to remain highly volatile due to higher exposure to exogenous shocks and faulty policies and structural issues. In other words, African countries, in general, has a high degree of permanence in unstable Regime. High volatility is observed to be a frequent occurrence in most economies in Africa as 26 of the 34 SSA economies had at least one switch-a ratio of 76.5%. As Byrne (2010) points out, one reason for this may be the reliance of most sub-Saharan African economies on primary commodity or extractive industries.
From Table 6 trade increases the probability of experiencing stable growth as well as FDI is found to encourage the probability of remaining in the stable regime. Political regime has a statistically significant impact on remaining in state 1, and its estimated coefficient is (as expected) negative. This implies that a change in the political regime from transitional to autocracy and from autocracy to democracy increases the probability for each country to maintain a sustainable growth rate. Indeed, as Hirschman (1958) emphasizes, with the absence of stable institutions, primary commodity and foreign owned extractive industries experience what he termed "enclave" type of development caused by the ability of primary products from mines, wells and plantations to slip out of a country without leaving much of a trace on economic growth. Also economies that transit from sustained growth to volatile growth are observed to have an increase in blackmarkets, racketeering and underground economy activities which deliberately underestimates GDP, creates shortages, leads to loss of revenue for governments and losses of legitimate industry thereby retard growth. Furthermore, by looking simultaneously at the evolution of the growth rate, the changes in the political regime and the regime change, it is clear that for countries experiencing frequent changes in the shift from one regime to another, there is a link between moving from dictatorship and political instability to democracy, and moving from State 2 to State 1. Interestingly, most economies which remain in state 1 are observed to autocratic regimes insinuating that Africa's "home grown" democracy might not be the best political system for sustainable growth in Africa. Moreover, as displayed in Table 4, this is not always substantiated as countries with the longest duration of remaining in a state 1 are democratic countries. Countries belonging to State 1 have almost the same probability of experiencing autocracy as democracy. However, most of the democratic countries belong to Regime 1, i.e. 19% of the democratic countries of our sample belong to Regime 2. Additionally, countries characterized by an unstable regime, like civil war and transitional regime belong as posterior classification to State 2.
On the other hand, instability in the growth pattern could be due to appreciation of the exchange rate, disinvestment from foreign investors and a lower lever of trade. Political regime has not a statistically significant impact on the transition from one State to another, thus shifting from democracy to autocracy or transitional regime and vice versa, do not affect the probability for a country to belong to move to the unstable growth rate group.

As the Table 7 highlights, most of the African countries have experienced changes from high volatility to low volatility and vice versa. The only exception, i.e. countries that do not pass through phases of high volatility to phases of low volatility are Chad, Djibouti, Gambia, Ghana, Ivory Coast, Mauritius, Sierra Leone and South Africa. Mauritius and South Africa have the distinction of being the only two African countries to have successfully transited to export led growth in manufacturing and services due to their strong institutions. Economies with the longest duration(span) of remaining in state 1 include Botswana,Gambia,Ghana,Ivory Coast,Mauritius and South Africa. With the notable exception of Botswana, none of these economies spent any period of time in state 2,thus they are characterized by high persistence levels. Botswana has had the highest average economic growth rate in the world since independence,averaging 9% per year. However,its economy stagnated in early 2000 due to a fall in diamond price(its primary export),but later grew at 6-7% target in subsequent years. Interestingly,it is rated as Africa's least corrupt nation by transparency international. Gambia, Ivory Coast, Chad, Djibouti and Ghana have experienced sustained slow growth over time, and as a result have been able to avoid output volatility. Benin, Burundi, Congo, Guinea, Mozambique, Sudan and Tanzania which do not change their economic behaviour over time all belong to $m = 1$, i.e. the low volatile with the highest mean of growth rate in the group. A common feature among these countries is the existence of long term authoritarian regimes. Recent studies surprisingly show positive correlation between authoritarianism and sustainable economic growth in developing and emerging economies particularly in Asia,with China,Singapore and Vietnam as prime examples. Lesotho,Senegal and Malawi experienced the highest number of transitions suggesting relatively quick recovery from shocks, while Botswana,Eq.Guinea,Mauritania and Zimbabwe experienced only high volatile growth rates within the period of observation. Economies with the longest duration in state 2 include Burundi, Chad, Ethiopia, Mauritania, Rwanda, Sudan, Burkina Faso and Niger. All aforementioned countries have been afflicted by civil wars and social unrest for the greater part of its history, and all countries are affected by their geographical location.
To conclude, the results show that though most Sub-Saharan countries have made significant breakthrough in economic growth in the past few decades, most countries find it difficult to maintain sustainable positive growth.

4 Conclusion

This paper applies a two state Markov Switching Model with time varying transition probability to 37 African countries over the period 1987–2011. We rely on the idea that the country's growth path is the result of different growth regimes, that could occur without obvious changes in country-fundamentals. Thus, we study the change of growth rate and its determinants, by directly modelling the transition probability as dependent on a set of explanatory variables, such as trade, FDI, exchange rate variation and political regime.
Our empirical results display the existence of two different growth regimes: a stable and an unstable growth regime. Furthermore, the probability of switching to the highly volatile regime negatively depends on exchange rate depreciation, trade and foreign direct investment, while democracy increases the probability of remaining in the stable growth regime.
Of course, a possible extension could be directly modelling the volatility of the growth rate in the estimation procedure, by estimating the variance of the growth rate itself in a bidimensional framework. Indeed, even if the regime varying variance allows for accounting for different output volatility, understanding the output volatility determinants, formalizing them also as

country-specific, could increase the possible policy implication derived from the model. Yet, other distributions, to account for the heavy tails growth rate distribution could be considered to implement the estimation procedure.

References

[1] Akinbobola, T. and O. Oyetayo (2010): "Econometric analysis of real exchange rate and domestic output growth in nigeria.", *International Journal of Academic Research*, Vol. 2, No. 5.

[2] Asiedu, E. (2002): "On the determinants of foreign direct investment to developing countries: is africa different?", *World development*, Vol. 30, No. 1, pp. 107–119.

[3] Becker, T., T. I. Becker, P. Mauro et al. (2006): *Output drops and the shocks that matter*: International Monetary Fund.

[4] Bilgili, F., N. S. H. Tülüce, and İ. Doğan (2012): "The determinants of fdi in turkey: A markov regime-switching approach", *Economic Modelling*, Vol. 29, No. 4, pp. 1161–1169.

[5] Blattman, C., J. Hwang, and J. G. Williamson (2004): "The impact of the terms of trade on economic development in the periphery, 1870-1939: Volatility and secular change",Technical report, National Bureau of Economic Research.

[6] Byrne, J. (2010): "Output collapse, growth and volatility in sub-saharan africa: a regime-switching approach", *The Economic and Social Review*, Vol. 41, No. 1, pp. 21–41.

[7] Cuaresma, J. C. and J. Wörz (2005): "On export composition and growth", *Review of World Economics*, Vol. 141, No. 1, pp. 33–49.

[8] Diebold, F., J. Lee, and G. Weinbach (1994): "Regime switching with time-varying transition probabilities", *lung*, Vol. 1, p. 3Í.

[9] Ding, Z. (2012): "An implementation of markov regime switching model with time varying transition probabilities in matlab", *Available at SSRN 2083332*.

[10] Dupasquier, C. and P. N. Osakwe (2006): "Foreign direct investment in africa: Performance, challenges, and responsibilities", *Journal of Asian Economics*, Vol. 17, No. 2, pp. 241–260.

[11] Durlauf, S. N., P. A. Johnson, and J. R. Temple (2005): "Growth econometrics", *Handbook of economic growth*, Vol. 1, pp. 555–677.

[12] Easterly, W., M. Kremer, L. Pritchett, and L. H. Summers (1993): "Good policy or good luck?", *Journal of Monetary Economics*, Vol. 32, No. 3, pp. 459–483.

[13] Filardo, A. J. (1994): "Business-cycle phases and their transitional dynamics", *Journal of Business & Economic Statistics*, Vol. 12, No. 3, pp. 299–308.

[14] Filardo, A. J. and S. F. Gordon (1998): "Business cycle durations", *Journal of Econometrics*, Vol. 85, No. 1, pp. 99–123.

[15] Glaeser, E. L., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2004): "Do institutions cause growth?", *Journal of economic Growth*, Vol. 9, No. 3, pp. 271–303.

[16] Golley, J. and R. Tyers (2006): "China's growth to 2030: Demographic change and the labour supply constraint", *The Turning Point in Chinas Economic Development, Asia Pacific Press, Canberra.*

[17] Hamilton, J. D. (1989): "A new approach to the economic analysis of nonstationary time series and the business cycle", *Econometrica: Journal of the Econometric Society*, pp. 357–384.

[18] Hausmann, R., L. Pritchett, and D. Rodrik (2005): "Growth accelerations", *Journal of Economic Growth*, Vol. 10, No. 4, pp. 303–329.

[19] Hirschman, A. O. (1958): *The strategy of economic development*, Vol. 58: Yale University Press New Haven.

[20] Jerzmanowski, M. (2006): "Empirics of hills, plateaus, mountains and plains: A markov-switching approach to growth", *Journal of development economics*, Vol. 81, No. 2, pp. 357–385.

[21] Jones, B. F. and B. A. Olken (2008): "The anatomy of start-stop growth", *The Review of Economics and Statistics*, Vol. 90, No. 3, pp. 582–587.

[22] Kerekes, M. (2012): "Growth miracles and failures in a markov switching classification model of growth", *Journal of Development Economics*, Vol. 98, No. 2, pp. 167–177.

[23] Kim, C.-J. (2004): "Markov-switching models with endogenous explanatory variables", *Journal of Econometrics*, Vol. 122, No. 1, pp. 127–136.

[24] Kim, C.-J., J. Piger, and R. Startz (2008): "Estimation of markov regime-switching regression models with endogenous switching", *Journal of Econometrics*, Vol. 143, No. 2, pp. 263–273.

[25] Kobrin, S. J. (2005): "The determinants of liberalization of fdi policy in developing countries: a cross-sectional analysis, 1992-2001", *Transnational Corporations*, Vol. 14, No. 1, pp. 67–104.

[26] Kuan, C.-M. (2002): "Lecture on the markov switching model", *Institute of Economics Academia Sinica.*

[27] de Mello Jr, L. R. (1997): "Foreign direct investment in developing countries and growth: A selective survey", *The Journal of Development Studies*, Vol. 34, No. 1, pp. 1–34.

[28] Pritchett, L. (2000): "Understanding patterns of economic growth: Searching for hills among plateaus, mountains, and plains", *The World Bank Economic Review*, Vol. 14, No. 2, pp. 221–250.

[29] Rodrik, D. (1999): "Where did all the growth go? external shocks, social conflict, and growth collapses", *Journal of economic growth*, Vol. 4, No. 4, pp. 385–412.

[30] Thapa, N. B. (2002): "An econometric analysis of the impact of real effective exchange rate on economic activities in nepal", *Economic Review: Occasional Paper*, Vol. 14, pp. 17–36.

Table 1: Descriptive statistics of the dependent variable

Variable	**Mean**	**Max**	**Min**	**Std. Dev.**	**Skewness**	**Kurtosis**
Growth	4.141495	19.45	-9.02	4.109465	-.0269384	4.600421

Table 2: Summary statistics of the explanatory variables

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
FDI	3.08	5.245	-8.59	46.49
Trade	-11.098	20.617	-126.56	48.96
ΔExch_Rate	17.056	93.856	-951	931.96
n. obs.		649		
Variable	**Freq.**	**Perc.**	**Cum.**	
Pol.Regime=1	157	24.19	24.19	
Pol.Regime=2	464	71.49	95.69	
Pol.Regime=3	28	4.31	100.00	
n. obs.		649		

Table 3: Descriptive statistics of the dependent variable classified by regime

Variable	**Mean**	**Median**	**Std. Dev.**	**Max**	**Min**	**25 perc**	**75 perc**	**St.Dev/Mean**
Regime 1								
Growth	4.240035	4.145	2.596867	19.45	-3.7	1.26	5.72	0.6124635
Regime 2								
Growth	3.712569	3.47	5.492451	18.87	-9.02	-3.09	7.32	1.479421

Table 4: Summary statistics of the explanatory variables classified by regime

Variable	Mean	Std. Dev.	Min.	Max.
Regime 1				
FDI	2.735	3.702	-6.9	31.64
Trade	-8.002	17.739	-126.56	29.66
ΔExch_Rate	28.204	107.799	-115.79	931.96
Variable	**Freq.**	**Perc.**	**Cum.**	
Pol. Regime=1	127	44.40	44.40	
Pol. Regime=2	159	55.60	100.00	
Pol. Regime=3	0	0	0	
N		286		
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Regime 2				
FDI	2.708	5.4	-5.84	46.49
Trade	-12.244	19.818	-119.89	48.96
ΔExch_Rate	4.074	81.324	-951	403.94
Variable	**Freq.**	**Perc.**	**Cum.**	
Pol. Regime=1	12	4.74	4.74	
Pol. Regime=2	215	84.98	89.72	
Pol. Regime=3	26	10.28	100.00	
N		253		

Table 5: Results: two switching parameter

	Time Fixed		TVTP	
	Coef	SE	Coef	SE
Growth rate:				
$growth_{it-1}$ (Regime 1)	0.8403***	0.0473	0.5730 ***	0.050
$growth_{it-1}$ (Regime 2)	0.2565***	0.038	0.2217 ***	0.051
$\alpha_{S_t=1}$ (Regime 1)	0.8226***	0.2209	1.8751**	0.2242
$\alpha_{S_t=2}$ (Regime 2)	2.9752 ***	0.2557	3.1420***	0.3921
$\sigma^2_{S_t=1}$	1.0000***	0.2235	2.003 ***	0.3501
$\sigma^2_{S_t=2}$	12.1173***	1.9648	20.0138 ***	3.2086
$p_{1,1}$	0.84		0.9658	
$p_{2,2}$	0.95		0.9600	
Duration Regime 1	6.14		6.55	
Duration Regime 2	16.83		8.07	
Observations	686		686	
K	2		2	
ℓ	-1692.27		-1695.67	
AIC	3440.9		3427.3	
BIC	3512.5		3507.9	
n. parameters	16		18	

Significance level: $***: 0.1\%$ $**: 5\%$ $*: 10\%$
Notes: K number of components; ℓ, log-likelihood
AIC$= -2\ell(.) + d$
BIC$= -2\ell(.) + d\log(n)$
where d is the number of parameters and n is the sample size
$p_{11} = p(S_t = 1|S_{t-1} = 1, \mathbf{z}_t) = p(\mathbf{z}_t)$ and $p_{22} = p(S_t = 2|S_{t-1} = 2, \mathbf{z}_t) = q(\mathbf{z}_t)$

Table 6: Time varying Transition probability: estimates

	$p_{1,1}$		$p_{2,1}$	
	Coef	SE	Coef	SE
Intercept	4.4883**	1.780	-0.2708	0.750
FDI	0.5326**	0.235	-0.1337*	0.074
Trade	0.0880**	0.040	-0.0287**	0.012
Pol. Regime	-1.5160*	0.811	-0.6875	0.438
ΔExch_Rate	0.0098	0.006	0.0052*	0.002

Significance level: $***: 0.1\%$ $**: 5\%$ $*: 10\%$

Notes: $p_{11} = p(S_t = 1|S_{t-1} = 1, \mathbf{z}_t) = p(\mathbf{z}_t)$ and $p_{22} = p(S_t = 2|S_{t-1} = 2, \mathbf{z}_t) = q(\mathbf{z}_t)$ where $\mathbf{z}_t$ is the matrix containing the explanatory variables affecting the transition probability.

Table 7: Countries' switching occasion

country	n. switching	switch to m=1	switch to m=2	n. obs in m=1	n. obs in m=2
Angola	2	1	1	1	3
Benin	1	1	0	14	1
Botswana	1	0	1	20	3
Burkina Faso	2	1	1	1	22
Burundi	2	1	1	1	22
Cameroon	3	2	1	16	7
Chad	0	0	0	0	19
DR Congo	1	1	0	4	4
Djibouti	0	0	0	4	0
Eq. Guinea	1	0	1	3	6
Ethiopia	3	2	1	2	14
G.-Bissau	6	3	3	4	6
Gabon	4	2	2	5	11
Gambia	0	0	0	17	0
Ghana	0	0	0	23	0
Guinea	1	1	0	2	2
Ivory Coast	0	0	0	20	0
Lesotho	12	6	6	9	11
Liberia	3	2	1	2	2
Madagascar	3	1	2	15	4
Malawi	5	3	2	9	9
Mali	4	2	2	11	9
Mauritania	1	0	1	2	20
Mauritius	0	0	0	23	0
Mozambique	1	1	0	6	15
Namibia	2	1	1	5	1
Niger	4	2	2	3	20
Rwanda	2	1	1	3	17
Senegal	5	3	2	13	10
Sierra Leone	0	0	0	2	0
South Africa	0	0	0	23	0
Sudan	2	1	1	2	15
Swaziland	3	2	1	15	7
Tanzania	1	1	0	18	2
Togo	2	1	1	4	16
Zambia	4	2	2	12	7
Zimbabwe	1	0	1	4	9
Total	82	44	38	318	294

Figure 1: QQ Plot

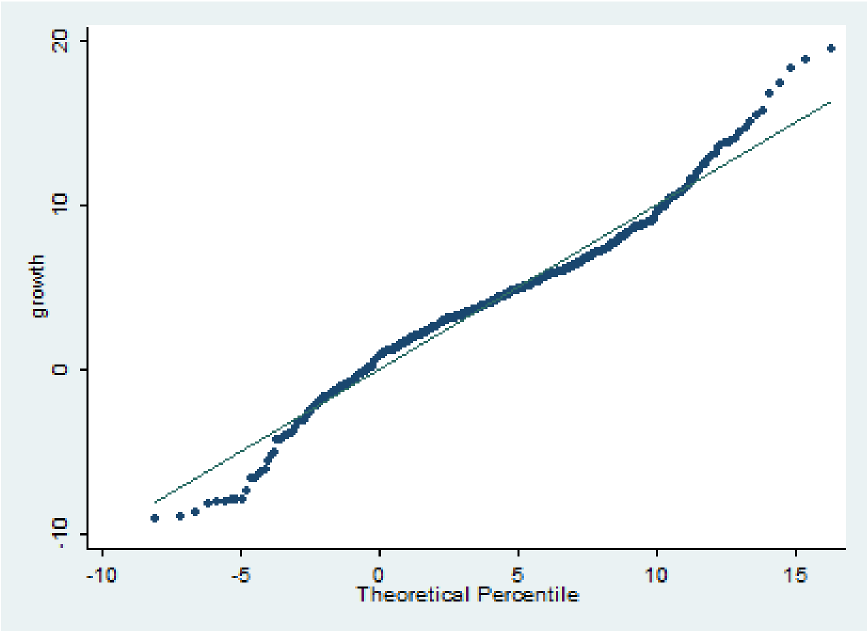


Figure 2: Fitting predicted values

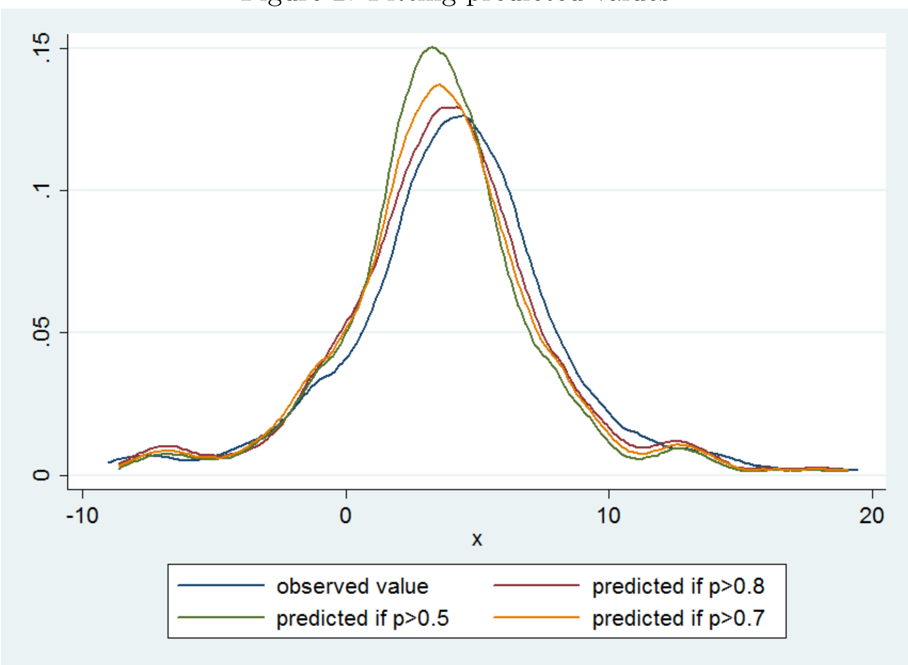
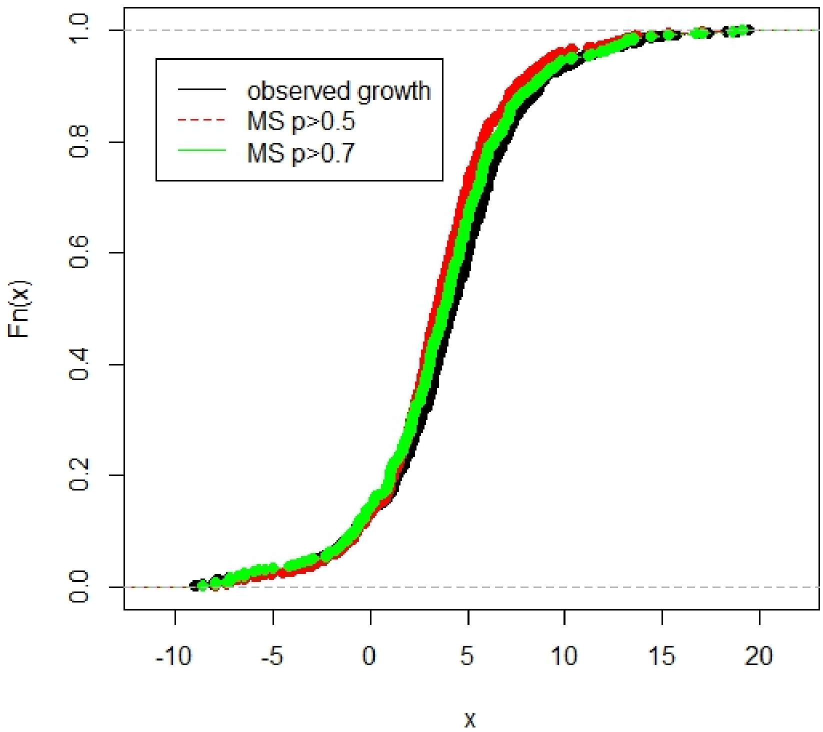


Figure 3: Empirical cumulative distribution function



Chapter 3

The determinants of country corruption. Unobserved heterogeneity and individual choice. An empirical application with Finite Mixture Models

Alessandra Marcelletti*†

Abstract

Corruption in public offices is found to be the reflection of country-specific features, however, the exact magnitude and the statistical significance of its determinants effect has not yet been identified. The paper aims to propose an estimation method to measure the impact of country fundamentals on corruption, showing that covariates could differently affect the extent of corruption across countries. Thus, we exploit a model able to take into account different factors affecting the incentive to ask or to be asked for a bribe, coherently with the use of the Corruption Perception Index. We assume that discordant results achieved in literature may be explained by omitted hidden factors affecting the agents' decision process. Moreover, assuming homogeneous covariates effect may lead to unreliable conclusions since the country-specific environment is not accounted for. We apply a Finite Mixture Model with concomitant variables to 129 countries from 1995 to 2006, accounting for the impact of the initial conditions in the socio-economic structure on the corruption patterns. Our findings confirm the hypothesis of the decision process of accepting or asking for a bribe varies with specific country fundamental features.

JEL Classification C14, C23, C29, D73
Keywords: Corruption, Finite Mixture Models, Concomitant Variables, Countries Classification

1 Introduction

The paper aims to investigate the main drivers behind the literature's discordant results on the interaction between corruption and the political and economic country-specific environment (see among others, Acemoglu and Verdier, 2000; Braun and Di Tella, 2004; Fréchette, 2006; Friedman et al. 2000; Husted, 1999; La Porta et al., 1999; Svensson, 2005; Treisman, 2000).
The paper relies on the idea that these discordant results could be due to three empirical problems characterizing the estimation of corruption: the unobserved heterogeneity, the omitted variable bias and the error in variable bias. To simultaneously solve these empirical problems and to test for the cross-country determinants of the phenomenon, we apply a finite mixture model with concomitant variables to an unbalanced panel of 129 countries from 1995 to 2006.

*Dipartimento di Economia Diritto ed Istituzioni, Università di Roma "Tor Vergata" via Columbia 2 00133 (RM) Italy, e-mail: marcelletti@economia.uniroma2.it tel:+39-0672595649"

†Dipartimento di Scienze Politiche, Luiss Guido Carli, Viale Romania 32, 00197, Roma, email: amarcelletti@luiss.it"

Concomitant variables allows for partially adjusting for the reverse causality between corruption and country-specific socio-economic structure. In fact, regarding to the latter problem, it is worth noting that there could be a reverse causality effect between socio-economic environment and corruption. In particular, the decision of undertaking corrupt activities could be affected by country's specific features - such as the legal, political and economic structure, as well as cultural and religious settings (see among others, Svensson, 2005) - as well as could affect the legal, political and economic environment.
To address the three statistical issues, firstly, we pose a random coefficient for all variables in the empirical design (Aitkin, 1999). It adjusts the parameters' estimation for the unobserved heterogeneity, that could be due to the differences between the country economic, political and social characteristics. Thus, the country invariant assumption on corruption determinants is relaxed, in favour of the existence of different effects on corruption in environments characterized by different country-specific fundamentals (quality of institutions, empowerment rights, economic growth, public expenditure and so on). In fact, as a by-product of the model, we cluster countries sharing the same effect of the unobserved (latent) variables. This means that groups are formed by countries having the same socio-economic structure, given the observed and unobserved covariates, and the prior probabilities conditioned on initial measures (at year 1995) of per capita GDP, fiscal rate and schooling.
Secondly, since corruption is a "per se" hidden action involving individual decision process (Banerjee et al., 2012), we assume that "corruption occurs at the interface of private and public sector"[1] (Ackerman, 1997), as result of individual rational decision process based on connected expected costs and benefits, as well as on individual subjective factors. Although the latter factors, which may include individual tastes and preferences, sense of justice, and attitude towards risk in committing illegal acts, are powerful in explaining corruption, they are generally unobservable and/or unmeasurable. Since their omission could lead to biased estimation, we include them as latent factors. Despite the fact that corruption perception indexes could generate empirical distortion, due to the difficulty of capturing the effective corruption level within countries (see among others Olken, 2007, and Donchev and Ujhelyi, 2007)[2], following Husted (1999) and Lancaster and Montinola (2001), we believe that the use of this kind of index, as like as the here used Corruption Perception Index, allows us to capture at least the likelihood of having corrupt agents within countries, according to how the public sector is seen to be. For this reason, we assume that agents undertake corrupt activities according to the perceived level of risk of punishment and perceived expected benefit deriving from corruption.
Entering into details, we model the intercept term and the slope parameters as dependent on hidden (latent) subjective factors, by imposing a latent structure for the covariates (Aitkin and Rocci, 2001). It is worth noting that in this way, the unobservable corruption determinants - sense of justice, cultural variables, individual attitude towards risk and so on - directly affect the individual perception about the risk of punishment and the economics incentive given by the environment.
Our contribution to the existing literature is that we exploit the finite mixture approach specificity to show that the controversial empirical results could derive from the omission of hidden subjective factors underlying the agents' decision process.
Our findings confirm that the decision of being corrupt could be influenced by political, economic and social country-specific characteristics, and they provide evidence for the presence

[1] In this way, it is clear that we drop out the possible corrupt behaviours between public officials (to gain position) or between private firms (to obtain procurement).

[2] Olken (2007) and Donchev and Ujhelyi (2007) empirically test the hypothesis that respondents' beliefs about corruption may not reflect the real level of corruption, generating bias empirical results.

of unobserved heterogeneity.
The paper is divided into fifth sections (including introduction and conclusions). The second section reviews the existence literature. The third section presents the empirical framework, firstly it explores the possible determinants of country corruption, secondly it illustrates the empirical model, finally, discusses the data. The forth section presents the results of the proposed approach, providing also a robustness check. The fifth section concludes.

2 Literature review

Since 1994, literature about corruption has expanded rapidly following the provision, among others, of the Corruption Perception Index and the Control of Corruption (respectively provided by the Transparency Index and the World Bank). This allows the researchers to empirically test the interaction between corruption and political, social and economic factors. Despite the availability of quantitative data on the phenomenon, its link with country fundamentals is debatable. As stressed in Table 1, empirical and theoretical findings agree that corruption interplays with political, social and economic environments, but they disagree on magnitude, sign and statistically significant of the relation.

Insert Table 1 about here

GDP (both in terms of growth rate and wealth per capita) is found and thought to be one of the main economic variable affecting the level of corruption within countries, even if its impact is debatable. Although it is found to reduce economic growth, by lowering private investments and Foreign Direct Investment (FDI) and distorting public services provision (Mauro, 1995; Anderson and Marcouiller, 2002; Wei, 1997), Paldam (2002) stresses that the correct causality relation is the reverse. In fact, in a transition model, he finds that corruption is a characteristic of poor and middle income countries, that disappears when they go through the grand transition to become high-income countries. Furthermore, many economists (Husted, 1999; Serra, 2004; Ata and Arvas, 2011; Svensonn, 2005) empirically test this result, by positively relating the level of GDP per capita and the growth rate of GDP with corruption. As theoretical explanation, they argue that a high level of GDP is associated with a high amount of Government resources, that can in turn be used in fighting corruption. In addition, rich and developed countries create a demand for institutional change and good government, that decrease officials' corrupt activities (Svensonn, 2005). Braun and Di Tella (2004) and Fréchette (2006), by using panel data, deviate from this commonly accepted result, by noticing that, because corruption has a pro-cyclical nature, 'moral standard are lowered during booms, as greed becomes the dominant force for economic decisions' (Braun and Di Tella, 2004, p.93). As stressed in Table (1), the contrasting results mainly concern the impact of the degree of intervention of the State in the economics and political environment (i.e. Government intervention and size) and the effect of the monitoring activity on corruption in the public sector (i.e. the extent of competition). Regarding to the latter, despite the fact that empirical estimation proves that competition, openness to trade and FDI are commonly linked to a low level of corruption (see among others Ades and Di Tella, 1996; 1999[3], and Robertson and Watson,

[3] Ades and Di Tella (1999) by using country's openness as an indicator of competition, empirically prove that economic competition decreases the extent of corruption; they also prove that the entrance of foreign investments corresponds to import competition and to reduce the rents for domestic firms and thus the rewards from corrupt activities.

2004[4]), theoretical results do not lead to a unique conclusion. Lambsdorff (1999) notices that a high level of competition lowers the rents of economic activities and the motive of public officer to seize parts of these rents. Ades and Di Tella (1999) highlight that the competitive pressure does not leave to the firms excess profit to pay bribes. On the other hand, Bliss and Di Tella (1997) develop a model in which the official, by inducing exit from the market, create the excess profit from which pay a bribe. Although the impact of competition among firms seems to have not a unique interpretation, it is commonly accepted that competition among public officials, by decreasing their monopoly power, reduces the propensity to accept a bribe (Ackerman, 1997). In fact, as Ackerman points out, the structural characteristics of the Government affect the demand for corrupted service. In turn, these factors together with the political features (including democracy, decentralization and unitarism) determine the quality of the Institutions.

Nonetheless, it is commonly accepted that a low level of corruption is associated with Institutions able to promote social cohesion, protect property rights as well as freedom of belief and religion, and ensure compliance with the law. Researchers directly measure the Institution quality by looking at the risk of expropriation within countries. Despite the lack of quantitative data hinders estimations, they agree that the lower the risk of expropriation (the higher the quality of institution), the lower the propensity to be asked for a bribe (Mocan, 2008).

As hinted before, the impact on corruption of the size of the State has not been reached a consensus in literature. Indeed, empirical estimation performed on the effect of the public expenditure in final goods as share of GDP - as a proxy for the Government size - leads to contrasting results. Bilger and Goel (2009), by using a quantile regression, and Adserá et al. (2003), find a negative relation between government size and corruption. On the other hand, starting from the idea that the State intervention and public spending give rise to rent-seeking, Goel and Nelson (1998) and Fisman and Gatti (2002), by using the number of public officials convicted for abuse in public office in USA, find a strong positive influence of government and local expenditure on corruption. The impact of the role of the State on corrupt phenomenon is debatable, even by looking at the Government intervention in the form of regulation and taxation. Since the intervention of the State in the market could generate partner advantages over rival, Treisman (2000) proves that the two variable are positively related [5]. Conversely, Friedman et al, (2000) conclude that a high degree of tax rate is associated with less unofficial activities, because of the stronger legal environment[6] Although democracy is thought to reduce the diffusion and the existence of the phenomenon, empirical findings do not completely confirm this theoretical result. Treisman (2000), among others, shows that only after 40 years, an uninterrupted democracy has a decreasing effect on the level of corruption, in terms of risk of being a victim of bribery. Moreover, democratic elections could create room for corruption. As Kunicova and Rose-Ackerman (2005) and Persson and Tabellini (2003) argue, party lists could represent an aspect of the democratic election that generates corruption; in fact, in their view, if there is not a direct link between voters and politicians, the latter agents could be less accountable by citizens.

Literature agrees about the relation between corruption and others government specific characteristics, as the legal origin of the country and the law system. In fact, British legal origin,

[4]Robertson and Watson (2004) analyze the rate of FDI's inflow. After controlling for cultural variables (gender, religion and so on), they prove that the more rapid the increase or the decrease in FDI into a country, the higher the perceived level of corruption.

[5]Treisman (2000) demonstrates that the State intervention is associated in 1996 with higher corruption, even if this is not significant either in the 1997 or the 1998 data.

[6]Friedman et al, (2000) also argued that the results depend on how the tax system is administered.

putting the attention on individual's right (private and property), is found to face a low level of corruption (David and Brierley, 1978, Finer, 1997, La Porta et al., 1998). French or Scandinavian legal origin, characterized by a greater attention to the power of the State, face a high level of corruption (Mocan, 2008). Similarly, whereas a common law system, developed in defense of property right and parliamentarianism, is found to lower corruption, the civil law system, concentrated on the sovereignty of the State, is found to increase corruption (David and Brierly, 1985; La Porta et al., 1999; Treisman, 2000). Gerring and Thacker (2004) confirm these results finding that parliamentarianism and unitarianism, by centralizing the political power and reducing the number of potential veto points, decrease corruption level.
Since corruption is a human activity, many economists show how cultural variables, as religion (e.g. La Porta et al., 1997; Treisman, 2000)[7] and education, can affect the propensity to behave illicitly. Regarding to the latter, it is worth noting that the stock of human capital, by interacting with institutional factor and by increasing citizens' monitoring ability, could play an important role in discouraging officials' corrupt activities. In fact, educated citizens have tools both to distinguish between corrupt and honest politicians behavior (Eicher, Penalosa and Ypersele, 2007) and to punish government abuses (Glaeser and Saks, 2006), as well as, once recruited, to improve efficiency of courts and Institutions (Svensonn, 2005); as a consequences, educated citizens discourages the extent of corruption, by increasing the institution ability in frightening corruption.
Nevertheless, despite the consensus reached about the relation between cultural variables and corruption, it is worth noting that, according to our knowledge, literature in this field does not take into account that the choice of undertaking a corrupt activity, as result of individual decision process, is influenced by unobservable and/or no-measurable factors, as the private sense of justice, and the attitude toward risk. In fact, we believe that the lack of these hidden variables makes literature about corruption characterized by the contrasting results showed in this section and summarized in Table 1.

3 Empirical Framework

3.1 The definition of corruption

Following Ackerman (1997), we define corruption as a phenomenon that "occurs at the interface of the public and private sector" [8]. The private (corrupter) agent acts as a *demander* of corruption, and he involves the misuse of public power position in order to obtain a legal or an illegal act, as well as an omission. The public (corrupt) agent acts as a *supplier* of the abuse and gains financial, material or non-material benefits, such as bribe or other type of compensation.
As rational agents, individuals accept or demand a bribe (or non-material reward) if it promises the greatest economic return, on the basis of Government's capability in deterring and punishing - the so-called *risk of punishment* (see among others, Becker, 1968) - the *economics*

[7]Researchers agree that hierarchical religions are positively related to the corruption level (e.g. La Porta et al., 1997). More into detail, Treisman (2000) finds that the larger the diffusion of the Protestantism in a country's population as of 1980, the lower the corruption perceived to be. In fact, according to him, Protestantism, as a traditional religion, is characterized by an independent church that can play a role in monitoring the state officials' abuse, conversing to hierarchies religion, characterized by interconnected state and church

[8]In this way we limit our model in the assumption that the phenomenon results only from the interplay between the private and the public agent, by dropping out the possibility of capturing criminal interaction among private firms (i.e., to obtain procurement) or among public officer (i.e., to obtain position)

incentives given by the environment in undertaking a corrupt activity (i.e. bribe and non material rewards) as well as subjective factors underlying the utility function. The latter factors include the individual tastes and preferences, the propensity of committing illegal acts, the sense of justice, as well as the attitude towards risk. Although they are powerful in explaining corruption, they are generally hidden and/or unmeasurable. Thus, a suitable model is obtained once it includes these subjective factors.

Additionally, since, by nature, corruption is an illicit activity, any data measuring level of corruption will suffer from measurement error. For example, data on the number of person convicted or persecuted could be unavailable and/or generate error-in-variable problems because of the country-specific nature of the national law. Following Husted (1999) and Lancaster and Montinola (2001), we address this problem by using the corruption perception index. We rely on the idea that the corruption perception could reflect the likelihood of having corrupt agents within country in a certain time window, by looking at how the public sector is seen to be. Furthermore, as Lambsdorff (2006) argues, the composite index nature and the fact that the index derives from international surveys data make this index comparable among countries. Thus, the expected costs - the probability of being getting caught, and any disutility regarding immorality (McChesney, 2010) - as well as the expected benefits - bribe, any type of compensation, and the possibility of avoiding bureaucratic system - are analyzed under the individual perspective. To do that, we formally include the unobserved and/or unmeasurable subjective factors that could differently affect the impact of the risk of punishment and the economics incentive on the spread of corruption across country.

The perceived *risk of punishment* rp_{it} for each country i at time t is determined by *macro* and *micro* level factors. As macro measure we consider the general model of crime, as the intervention of the State in the economic system in terms of size of the State[9], gs_{it}, accounting for the role of democracy, d_{it}, and the Government fractionalization, ps_{it}.

In order to have a comprehensive measure of the perceived capability of the State in increasing the likelihood of being getting caught[10], the perceived level of protection of civil and political rights, r_{it}, and the degree of independence of judiciary system, j_{it} enter the model as *micro* level factor. Thus, the risk of punishment function reads as:

$$rp_{it} = f(r_{it}, gs_{it}, ps_{it}, d_{it}, j_{it}) \tag{1}$$

The temptation of behaving illegally depends on the size and the role of the State - as *macro* measure - and on the private perception about the effective intervention of the State in case of non compliance with the law - as *micro* factor. In fact, it seems reasonable to believe that a country characterized by an independent judiciary system and a high degree of protection of civil and political rights is viewed by citizens as active in ensuring compliance with the law, as well as in procuring the necessary tools to the people to participate in checking and denouncing corruption in public office.

The expected returns are determined by the *economics incentive* given by the country-specific environment. Thus, the individual's gain depends on *direct* - bribe and any type of compensation - and *indirect* returns - the avoidance of the bureaucratic system. In other words, the expected benefit from corruption is given by the possible gain for the private agent of avoiding the bureaucratic system, as well as of obtaining a reduction in the tax burden. In this way, the

[9]Following the economic literature (see among others Kotera, Okada and Samreth, 2012) the Government size is proxied by the general government final consumption expenditures, as a share of GDP.

[10]In the economics of crime, the risk of punishment is estimated by using objective data, as the per capita police expenditure, on the deterrent effect that police presence may have on the criminal behavior.

probability to ask for or accept a bribe is affected by the impact on the private agent's activity of the bureaucratic service b_{it}, accounting for the country-specific financial environment in terms of State intervention in financial services and openness, f_{it}, as well as openness to trade. In fact, as discussed in Section 2 (see among others, Ades and Di Tella, 1999) openness to trade [11] o_i, seems to affect corruption level within countries. More into details, following the economics literature, we believe that the international commercial activity could decrease the monopolistic rents enjoyed by bribe. Thus, openness to trade, as a proxy, enters the model. Summing up, the *economic incentives* ei_{it} may be modelled as:

$$ei_{it} = f(b_{it}, f_{it}, o_{it}) \tag{2}$$

To account for the impact of country-specific socio-economic structure on corruption, we add as concomitant variables the initial level of education, $educ_{i,95}$, of the GDP per capita, $gdp_{i,95}$, as well as the tax burden, $t_{i,95}$.

3.2 Empirical Model Specification

To address the unobserved heterogeneity issue, we apply a finite mixture model with concomitant variables. By allowing a latent variable to enter the estimation process, the resulting parameters account for unobservable and unmeasurable factors. In this way, we can capture the impact of the different socio-economic country-specific environment on the extent of corruption. On the basis of the latent variable, the entire sample n is clustered in $k = (1, \ldots, K)$ subgroups, in which the country invariant assumption about the effects of the corruption determinants holds. Furthermore, the relationship between behavioral and socio-economic variables and corruption is captured by allowing the weights of the mixture (i.e., the group size) to depend on concomitant variables, i.e. variables affecting the country probability of belonging to a specific cluster.
Formally, let y_{it} be the vector containing the realized conditionally independent and identically distributed random variables of the outcome recorded for each country $i = (1, \ldots, n)$ at time $t = (1, \ldots, T)$. Let $\mathrm{x}^1_{it} = rp_{it}$ be the set of covariates capturing the *risk of punishment*, and x^2_{it} the *economics incentive* in undertaking a corrupt activity (as respectively described in equation (1) and (2)). We further denote u_i the set of unobserved random factors, that accounts for country-specific heterogeneity and dependence among covariates. These variables include subjective factors (as cultural traits, attitude towards risk and so on), underlining the choice of being corrupt.
The need of using this statistical approach is justified also by the fact that when we estimate a model with these omitted variables we could have bias in the estimated parameters and in their related significance. For example, the estimation of corruption determinants through an OLS-based approach could conduct to bias results since OLS assumes that on average the effects on y_{it} of the matrix x^1_{it} and x^2_{it} are homogeneous among countries (see among others, Durlauf et al., 2005). In particular, considering a standard linear model, where $\boldsymbol{\beta} = \{\boldsymbol{\beta}^{\mathsf{T}}{}_0, \boldsymbol{\beta}^{\mathsf{T}}{}_1, \boldsymbol{\beta}^{\mathsf{T}}{}_2\}$ denote the parameters vector, we can write the regression function as:

$$E(\mathrm{y}_i|\mathrm{x}^1_{it}, \mathrm{x}^2_{it}) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}^{\mathsf{T}}_1 \mathrm{x}^1_{it} + \boldsymbol{\beta}^{\mathsf{T}}_2 \mathrm{x}^2_{it} \tag{3}$$

where parameters in $\boldsymbol{\beta}$ result to be equivalent in terms of sign and magnitude among different countries, even if each country could be characterized by different (measurable or omitted

[11] As trade openness measure we use the sum of the share of imports and export dividing by the GDP.

and/or unmeasurable) economics, social and institutional features.
To avoid the homogeneity assumption among the entire sample, we formalize the objective regression function as:

$$E(\mathrm{y}_i|\mathrm{x}^1_{it}, \mathrm{x}^2_{it}) = \boldsymbol{\beta}_{0i} + \boldsymbol{\beta}^{\mathsf{T}}_{i1}\mathrm{x}^1_{it} + \boldsymbol{\beta}^{\mathsf{T}}_{2i}\mathrm{x}^2_{it} \tag{4}$$

where all the parameters are jointly determined by the correlated latent traits u_i and the common shared effect $\boldsymbol{\beta}$, such that $\boldsymbol{\beta}_{1i} = \beta_1 + u_i$, $\boldsymbol{\beta}_{2i} = \beta_2 + u_i$, $\beta_{0i} = \beta_0 + u_i$. In other words, $\boldsymbol{\beta}_{0i}$ $\boldsymbol{\beta}_{1i}$ and $\boldsymbol{\beta}_{2i}$ are deviations from the common shared effects measured by, respectively, β_0 β_1 and β_2 and vary among countries in function of the latent covariates u_i, i.e., u_i is the unobserved heterogeneity characterizing the different socio-economic structure among countries, correlated for the different $\boldsymbol{\beta}$. Given the conditional independent assumption, by assuming that the dependent variable is drawn from the normal distribution, the mixture model becomes:

$$\begin{aligned} f_i &= f\left(\mathbf{y}_i \mid \mathbf{x}_i, \mathbf{u}_i\right) = \prod_{t=1}^{T} \left\{ f\left(y_{it} \mid x_{it}, u_i\right) \right\} = \qquad (5) \\ &= \prod_{t=1}^{T} f_{it} = \prod_{t=1}^{T} \left\{ \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2} \left(y_{it} - (\beta_0 + u_i) - x_{it1}(\beta_1 + u_i) - x_{it2}(\beta_2 + u_i) \right)^2 \right] \right\} \end{aligned}$$

Since the latent variables is unknown, we have to integrate it out. Moreover, in order to have no restrictive assumption on the distribution of $\mathbf{u}_i$, we leave $G(\cdot)$ completely unspecified, obtaining the following likelihood function:

$$L(\cdot) = \prod_{i=1}^{n} \left\{ \int_u f_i \mathrm{d}G(\mathbf{u}_i) \right\} \tag{6}$$

Furthermore, since we believe that the *prior* probability of belonging to a certain groups k is affected by the country-specific socio-economic structure, as the GDP per capita, the education level and the tax burden, we allow the weights of the mixture density to depend on these variables. Formally, let c the set of concomitant variables, such that $c = f(gdp_{i,95}, t_{i,95}, educ_{i,95})$, as discussed in Section 3.1, and α the associated parameter vector, such that we can rewrite the prior probability as $\pi_k = f(c, \alpha)$, where $\forall c \sum_{k=1}^{K} \pi(c, \alpha) = 1$ and $\pi(c, \alpha) > 0$. Following Dayton and Mcready (1988), a multinomial logit model is assumed for π_k, where the first component is the baseline (e.g., McLachlan and Peel, 2000) model to estimate the weights of the mixture:

$$\pi_k(c, \alpha) = \frac{e^{c^{\mathsf{T}}\alpha_k}}{\sum_1^K e^{c^{\mathsf{T}}\alpha_k}} \tag{7}$$

And the estimated posterior probability is

$$\hat{w_{ik}} = p(u_i = 1|\mathbf{y}_i) = \frac{\pi_k(c, \alpha) f\left(\mathbf{y}_i \left|\mathbf{x}_i, \mathbf{u}_k\right.\right)}{\sum_1^K \pi_k(c, \alpha) f_{ik}} \tag{8}$$

where $\mathbf{x}_i = \{x^1_{it}, x^2_{it}\}$ and for sake of simplicity we denotes with $f_i k = f(\mathbf{y}_i|\mathbf{x}_i, \mathbf{u}_k)$ the response distribution in the k-th component of the finite mixture. Thus, by approximating the integral in (6) as a sum on a finite number of locations K, the resulting likelihood function is:

$$L(\cdot) = \prod_{i=1}^{n} \left\{ \sum_{k=1}^{K} f\left(\mathbf{y}_i \left| \mathbf{x}_i, \mathbf{u}_k \right.\right) \pi_k(c, \alpha) \right\} = \prod_{i=1}^{n} \left\{ \sum_{k=1}^{K} [f_{ik} \pi_k(c, \alpha)] \right\} \tag{9}$$

In order to get the parameters' estimate, let $\boldsymbol{\delta} = (\boldsymbol{\beta}_{0k}, \boldsymbol{\beta}_{1k}, \boldsymbol{\beta}_{2k}, \alpha_k, \pi_1, \ldots, \pi_{k-1}, \mathbf{u}_k, \sigma^2_{u_k})$[12] the *complete* vector containing the unknown parameters of the model. Let K to be treated as fixed and estimated via penalized likelihood criteria in the parameters estimation process (McLachlan and Peel, 2000).
Nevertheless, since the label component indicators is missing, the EM algorithm naturally arises to get the ML estimation. Formally, let $\mathbf{y}_i^c$ the complete data vector containing the feature data and the unobservable component indicators $\mathbf{z}_i$, where $\mathbf{z}_i = (z_{i1}, \ldots, z_{ik})$ is the unobservable vector of component indicators, containing dummies variable z_{ik} equal to 1, if the observation i has been drawn from the k component of the mixture, and 0 otherwise. Thus the complete data likelihood reads:

$$L(\cdot) = \prod_{i=1}^{n} \prod_{k=1}^{K} \{\pi_k(c, \alpha) f(\mathbf{y}_i | \mathbf{x}_i, \mathbf{u}_k)\}^{z_{ik}} \tag{10}$$

and the corresponding complete data log-likelihood reads as follows:

$$\ell_c(\cdot) = \sum_{i=1}^{n} \sum_{k=1}^{K} \hat{z}_{ik} \left[\log(\pi_k(c, \alpha)) + \sum_i \log(f_{ik}) \right] \tag{11}$$

The complete parameters vector's estimates are simultaneously obtained by performing the EM algorithm[13]. For computational details see, among others, Dayton and Mcready (1988).
It is worth noting that the prior probability for each country of belonging to cluster k is computed conditional to the concomitant variables, where the parameter $\hat{\alpha}$ identifies whether the concomitant variables can affect the probability for each country of belonging to a certain cluster with respect to the benchmark group (group 1).

3.3 Data

Our work is based on an unbalanced panel for 129 countries from 1995 to 2006. Data are retrived from different database.
The dependent variable is the Corruption Perception Index (CPI) provided by Transparency International. It covers 130 countries from 1995, and it is constructed by aggregating various sources of survey data, at most 13. It ranges between 0 and 10 such that the highest the score, the lowest the extent of corruption.
Table 3 displays summary statistics for the Corruption Perception Index. Skewness and kurtosis, of 0.79 and 2.38 respectively, show a departure from the symmetric and flatness common measure of the normal distribution. To complement the non-normal distribution of the dependent variable, the Q-Q Plot and the Shapiro-Wilk test are performed. Both confirm departure from normality of the sample data. Particularly, the QQ plot (Figure 1) shows that data points fail to especially in the tails to follow the line very.

Insert Figure 1 about here

[12]Since the prior probabilities by definition sum up to 1, one of the mixing proportion is redundant
[13]Model parameters are estimated through the Flexmix package, as developed by Grünn and Leisch (2008).

Shapiro-Wilk and Anderson and Darling normality test (Shapiro and Wilk, 1965 and Anderson and Darling, 1954) show that the null hypothesis of normal distribution of the data could not be accepted (p-values for H_0 is $p - value = 2.2e^{-16}$ for both tests; A statistics for Anderson and Darling test is equal to 42.1073 and W statistics, for the Shapiro and Wilk test 0.8901). As argued in the Introduction and in Section 2, despite its widespread use, the CPI is affected by some criticisms. Nevertheless, corruption perception and actual level of corruption within countries are found to be closely related (see among others, Fisman and Miguel, 2007), so CPI could be considered an useful (even if not exhaustive) measure for corruption. For this reason, we employ CPI as dependent variable taking into account that our results concern the perceived level of corruption across countries. The independent variables are divided in two groups: *risk of punishment* determinants and *economics incentive* determinants. Variables belonging to the first set of covariates are:

- the government size, *gce*, measured as the general government final consumption expenditure as percentage of the GDP (World Development Indicators, $http://data.worldbank.org$);
- the democracy *demo*, as a dummy equal to 1 if the regime is democratic and 0 otherwise (Cheibub, Gandhi and Vreeland database $http://www.systemicpeace.org/polity/polity4.htm$);
- the Parliament power during the legislative iter, *ovs* measured as vote of opposition parties over total votes (Cheibub, Gandhi and Vreeland database, $http://www.systemicpeace.org/polity/po$
- the level of political rights and civil liberties, *status*, a categorical variable coded 1 if the protection is complete, 2 if it is partly, and 3 if it is absent ($http://www.freedomhouse.org$);
- the independence of the judiciary system, *jud*, as a dummy variable coded 1 if there is an independent judiciary, and 0 otherwise (Henisz political constraints index, $http://www - management.wharton.upenn.edu/henisz/$).

Variables belonging to the *economics incentive* set of covariate are:

- the freedom to start business activities in no-financial sector, *business* (Heritage Foundation);
- the government intervention in the financial service, and of the freedom of opening and operating financial services firms, $finance$ (Heritage Foundation);
- the *openness to trade* at constant price at 2005, *openk* (Penn World Table 8.0, $http://www.rug.nl/research/ggdc/data/penn - world - table$).

As concomitant variables we use:

- $fiscal_{95}$: the initial level of the composite index measuring the State fiscal policy over individuals and firms (Heritage Foundation);
- $educ_{95}$: the initial level of the average number of years of education of men, aged 25 and older (The Quality of Government Institute, University of Gothenburg);
- $rgdpch_{95}$: the initial level of per capita income (Penn World Table 8.0).

Data retrieved from the Heritage Foundation (*http : //www.heritage.org/index/download*) range between 0 and 100, where the highest the score, the highest the level of freedom.

Table 2 describes the data (as average of the index, or as frequency) on the basis of the level of corruption. Following the Transparency Index, we divide corruption in three categories, high, medium and low according to the index score. An index between 0-3 is associated to high corruption, 4-6 with medium corruption, and 7-10 to low corruption.

Insert Table 2 about here

Descriptive statistics seem to confirm that political setting as well as economics variables are strictly related to the extent of corruption. Even if the average of some country fundamentals (democracy, independent judiciary, education, as well as business and financial freedom) presents the expected relation with CPI, others variables have an uncommon relation with corruption. For example, the protection of political and civil rights is ensured more in countries with a medium corruption perception, both in terms of index average and frequency, rather than in low corrupted countries. Regarding to the most debated variables, GDP per capita, government size (in terms of public expenditure) and government intervention, data distribution shows that low corruption is more reasonable in the richest countries, characterized by the highest government size, and the lowest intervention of the State in the economics environment, even if the difference in the fiscal index is low among low and medium corrupt countries.

Before proceeding to analyze the estimation results, it is useful to have a look at the response variable distribution, in order to better understand that the need of clustering the entire sample in sub-populations is due also to the uncommon distribution of the CPI. In fact, by looking at the CPI density, it seems more reasonable to consider it as a mixture of a normal variables, rather than as a common normal distribution (see Figure 2).

Insert Figure 2 and Table 3 about here

4 Results

In this section we present results obtained by estimating the model from equation (4), using generalized least squares (GLS henceforth) and OLS with Fixed Effect (FE henceforth) approach, as parametric benchmark. Secondly, we show the parameters' estimate of the finite mixture model with concomitant variables.

Table 4 displays results for both GLS and OLS with FE.

Insert Table 4 about here

GLS results are consistent with the empirical findings obtained in literature about the impact of the main fundamentals on corruption (empowerment rights, financial and business index), whilst openness to trade is found to have not a statistically significant deterrent effect on corruption, as well as the tax burden. Moreover, regarding to the role of the State GLS estimation provides evidence in favour of Fisman and Gatti (2002) results, suggesting a negative correlation between Government size and corruption perception (it is worth reminding that the CPI index goes from 0, highest level of perceived corruption, to 10, absence of corruption). OLS with FE estimation presents statistically significant association among dependent and explanatory variables only for finance ($\hat{\beta} = 0.004$), democracy ($\hat{\beta} = -0.29$) and judiciary

independence indexes ($\hat{\beta} = 0.272$), and in each case the impact of these variables is lower than the one found through the GLS method.
In the next step, we present the results of the finite mixture approach with concomitant variables. After conditioning the weights of the mixture to depend on concomitant variables, data shows that our sample could be divided in six different groups. Table 5 shows that penalized criteria (BIC, AIC and ICL) values have been minimized with a discrete model log-likelihood with 6 components.

Insert Table 5 about here

Since the ratio between the number of observations assigned to the corresponding clusters and the one where the posterior probability is greater than N (with $n = 10^{-4}$) is around 0.7 in all sub-groups, but the first, we have evidence of medium well-separated components (see Table 6). This implies that there is not a significant overlap among components, but the first and the second (McLachlan and Peel, 2000).

Insert Table 6 about here

Thus, we can suppose the existence of distinct socio-economic rules characterizing countries in separated groups (at least in our sample).
Table 7 displays parameters' estimate from a finite mixture model with concomitant variables, showing that political, social and economic country-determinants vary their impact in terms of sign and magnitude on corruption perception among clusters.

Insert Table 7 about here

Indeed, the $\boldsymbol{\beta}_k$ parameters attached to both the *risk of punishment* and the *economics incentive* vary to capture the effect of the hidden factors, such as the sense of justice, the attitude towards risk, and so on, on the corruption determinants.
The most interesting result is the changing in sign of the parameters associated to the role of the Institution in fighting corruption, in terms of size and market intervention. As hinted before, the impact of the government expenditure on the expected cost in undertaking a corrupt activity depends on the individual perception about the strengthen of Institution in fighting corruption, that in turn depends on the individual perception about the role of the government in the economy. In fact, even if the structural characteristics of the Government are thought to affect demand for corrupt activities (Ackerman, 1998), the actual effect on corruption depends on how monitoring activity of the State is seen to be by agents. According to the literature debate, it is not surprising that in the first, fourth, fifth, and sixth component the government size has a positive (even if different in terms of magnitude) effect on corruption, confirming Bilger and Goel (2009) results, while in the second cluster it is negative, empirically proving that the increase in the public spending could give rise to rent-seeking (Fisman and Gatti, 2002). Indeed, despite the quite similar value of the government expenditure among class 2, 3 and 4 (see Table 9 for cluster composition and Table 10 summary statistics divided by cluster), the impact on corruption is highly different among the three sub-populations. Indeed, the $\hat{beta}$ parameter associated to the *gce* variable is significant in the second and fourth component, and it is respectively equal to -2.224 and 3.107, while is not statistically different from zero in the third group. This could be explained by the latent structure conditioning the parameters estimates. As it will be better explained in the following, since the country fundamentals

characterizing countries in class 4 are good (see Table 9 and Table 10), an increase in the government expenditure could not leave room for corruption.
In line with the idea that government intervention could "create room for corruption", as well as could deter it, it is worth noting that also the State intervention in the economics activity has discordant effect on corruption. Indeed, the financial freedom index is found to discourage the likelihood of asking for and/or accepting a bribe in all components, but the fourth. Class 3, which includes a mix of developed countries (see Table 9) with an index average of CPI and openness to trade lower than the one of class 4 but highest than the others (see Table 10), is characterized by a negative impact of the "financial freedom", (see Table 7), with a $\hat{beta}$ parameter associated to *finance* equal to -0.0049. We recall that this index is formed by two indistinct components: the degree of intervention of the State in the financial system, and the difficulty of opening and closing a financial service activities, also for foreign firms. Thus, we may conclude that the negative impact of *finance* on corruption (in the fourth component) could be due to presence of the State in the financial system, i.e. the highest the intervention of the State in the financial system, the highest the corruption perception in countries belonging to class 4.
To complement the analysis of government intervention, we account also for the role of democracy. Its $\hat{\beta}$ parameter is found to be significantly different from zero only in the first group (where the $\hat{beta}$ parameter associated to *demo* is equal to -0.5762). This confirms our idea that the strength of the State in fighting corruption changes across countries.
No discordant results, according to our knowledge, are achieved in literature on the effect of the independent judiciary system on corruption. For this reason, it is not surprising that the parameter associated to this variable has always the same effect in terms of sign and statistically significant in all clusters. This is due to the fact that the presence of an independent judiciary system, by increasing the perception about the capability of the State in ensuring compliance with the law, increases the perceived risk of punishment that in turn decreases corrupt acts.
At this point it is worth noting that despite class 5 and class 6 have a similar economics and political structure, only empowerment rights, judiciary independence index and government size affect corruption pattern in class 6 (business and finance freedom index's coefficient are significant but quite small), suggesting that in this cluster the role of the State in increasing the perceived compliance with the law is the unique variable affecting the corruption perception. On the other hand, estimated parameters in class 5 suggest that both the risk of punishment and the economics incentive in undertaking a corrupt activities significantly impact on corruption, even if the magnitude of the economics variable is low, but openness to trade.
In line with the literature debate, the perception on how the economics incentive could affect the extent of corruption is not unique. Our results confirm the idea that international competition decreases the extent of corruption. In fact, by increasing competition firms face a reduction in the extra profit used for paying bribe (see among others, Ades and Di Tella, 1999; Robertson and Watson, 2004 and Lambsbdorff, 1999); at the same time competition increases the monitoring activity played by other firms in the market, that in turn has a deterrent power on corrupted actions. Moreover, openness to trade is found here to have a deterrent role for the first, second and fifth component. In the other groups the impact on corruption is not statistically significant different from zero.
According to our theoretical framework (see Section 3.1), the will of avoiding the bureaucratic system is the main economics incentive discouraging incentive of undertaking a corrupt activity, since the parameter associated to the business index does not change its impact in terms of sign, but it is not statistically significantly different from zero in the second component.
In order to understand if the differences among countries in the corruption patterns could be

due also to the different starting points in the GDP per capita, the tax burden and the education quantity, we directly modeled prior probabilities of belonging to a certain cluster through concomitant variables.

Table 8 shows us the impact of the initial conditions on corruption patterns in different clusters, by taking the first category as benchmark. Our benchmark (group 1) contains medium corrupted countries[14] (CPI around 4.2, with standard error of 1.18), with a negative random term and the highest "fiscal freedom index" (i.e., the lowest tax burden).

Insert Table 8 about here

The starting condition that mainly affects the corruption perception is the GDP per capita measured in 1995. In fact, the wealth per capita increases the likelihood of belonging to the two "virtuous" groups (3 and 4) and to class 2 (containing medium income countries, with CPI around 5.0) relative to that of belonging to the benchmark group (class 1). On the other hand, the GDP per capita decreases the likelihood of belonging to the two poorest groups (class 5 and 6). Coherently with the economics literature, the initial level of GDP has different impact on corruption. In fact, the high level of GDP could be associated with a high amount of government resources and also to a high demand for institutional change and, thus, better government (Svensson, 2005). Moreover, the procyclical nature of corruption is the theoretical framework used to justify the negative effect on the extent of corruption, as it is derived for the sixth component (Braun and Di Tella, 2004). As Table 8 shows, the initial level of education in 1995 has a significant explanatory variable only in class 6, meaning that the quality of education raises the likelihood of belonging to class 6 (the highest corrupted) relative to that of belonging to the benchmark group. It is worth noting that even if class 6 is one of the poorest group (in terms of the GDP per capita) and with the lowest corruption perception index, the quality of education is relative high to that of class 1. Summing up, the economics structure is the main factor affecting the likelihood of belonging to "virtuous" groups.
Table 9 illustrates the country (posterior) classification.

Insert Table 9 about here

As showed in Table 9, countries' classification in clusters (by using the finite mixture approach) is sufficiently satisfactory, and the estimated classes show evidence of homogeneity within groups.
Indeed, by having a look at the clusters' composition, it seems clear that the "virtuous" countries, both in terms of corruption perception and country-fundamentals (see Table 10 for further detail on the independent variable summary statistics), are all clustered in class 4[15], in which the key country-characteristics affecting the corruption pattern are the business and the judiciary independence index, as well as is among others the government expenditure. Class 5[16] and 6[17], with both a statistically significant value of the random term, include the poorest

[14]Class 1 contains: Algeria, Botswana, Brazil, Burkina Faso, Cameroon, Chad, Cyprus, Estonia, Gabon, Greece, Hungary, Jordan, South Korea, Lebanon, Mali, Mauritius, Nepal, Niger, Peru, South Africa, Turkey

[15]Class 4 contains: Canada, Denmark, Finland, Germany, Ireland, Netherlands, Norway, Oman, Singapore, Sweden, United Kingdom, United States

[16]Benin, Cambodia, Dominican Republic, El Salvador, Fiji, Ghana, Guinea, India, Jamaica, Lithuania, Madagascar, Malawi, Mauritiana, Mexico, Morocco, Mozambique, Namibia, Poland, Senegal, Sri Lanka, Swaizland, Vietnam, Yemen, Zimbabwe

[17]Class 6 contains: Albania, Argentina, Armenia, Azerbaijan, Bangladesh, Bolivia, Bulgaria, Burundi, China, Congo, Croatia, Czech Republic, Ecuador, Egypt, Gambia, Georgia, Guyana, Honduras, Indonesia, Iran, Kenya,

countries, with the highest tax burden.
To conclude, our empirical findings show how the country-specific fundamentals affect in different way the extent of corruption among countries, according tot eh latent socio-economic structure. Thus, the incentives of asking for and/or being ask for a bribe change their impacts on the basis of the country-specific unobservable environment. Furthermore, despite parametric benchmark models deal with heteroscedasticity, heterogeneity and orthogonality among explanatory variables, they are not able to solve simultaneously the above mentioned empirical problems affecting the estimation procedure. In fact, by comparing the empirical density of the CPI, and the estimated density obtained by using the finite mixture and the GLS approach (see Figure 3), it is clear that the data generating process (for our sample) is better approximate by the mixture model, rather than the parametric approach.

Insert Figure 3 about here

The empirical density shows a strong evidence of the presence of heterogeneity in the sample, confirmed by the parameters estimates obtained by employing finite mixture model.

4.1 Robustness Check

Since the finite mixture model is a semi-parametric approach, standard parametric method to test the goodness of fit can not be applied. Firstly, following Aitkin (1997) and McLachlan and Peel (2000), we compare the empirical distribution function of the observed data (ECDF) with the fitted values derived both from the finite mixture and the GLS approach. The ECDF (Figure 5) shows that the data generating process is better approximated by the finite mixture approach instead of the GLS.

Insert Figure 5 about here

Secondly, to test the likelihood of having 6 non-overlapping subgroups, we follow a bootstrap based approach (Aitkin et al., 1981; Romano, 1988). In fact, since the regularity conditions do not hold for the log-likelihood statistic, we can not perform this test by using a simple likelihood ratio test (e.g., McLachlan, 1987). Thus, we test the following system of hypothesis, by an appropriate resampling[18], with say B (B=1000) number of replications:

$$\begin{cases} H_0 & k = j \\ H_1 & k > j \text{ or } k < j \text{ with } j = k \end{cases}$$

We test the *unimodal* distribution assumption by assuming $j = 1$ as in (4.1). Table 11 shows evidence of multimodal distribution (*p-value*=0.0102). To check for the compatibility between the estimated number of clusters and the data, we repeat the test as in (4.1) with $j = 6$. As Table 11 shows, we can state that having 6 sub-groups is reasonable according to the data.

Insert Table 11 about here

Laos, Latvia, Lesotho, Moldova, Mongolia, Nicaragua, Pakistan, Panama, Paraguay, Philippines, Romania, Russia, Rwanda, Slovakia, Syria, Tajikistan, Tanzania, Thailand, Uganda, Ukraine, Uzbekistan, Venezuela, Zambia.

[18] Following McLachlan (1987), the log-likelihood test statistic for the hypothesis test as in (4.1) can be bootstrapped as follows. Firstly, generate a bootstrap sample from the mixture density, where the $\hat{\boldsymbol{\delta}}$ is the ML estimator of the complete parameters vector formed under the null from the original sample. Then fit the mixture model $k = j$ and $k \leq j$ and compute the log-likelihood test statistic for the bootstrap sample. The process is repeated an independently number of times, say B equal to 1000.

5 Conclusion

To understand the literature debate about corruption determinants, the paper empirically tests the hypothesis that the individuals decision of being corrupt varies according to specific fundamentals, as well as unobservable and/or non-measurable variables (such as individual tastes and preferences, attitude toward risk, propensity of committing criminal acts), i.e. the heterogeneity is due to unobserved latent differences between the economics, political and social country-specific environments. To do that, we apply a finite mixture model with concomitant variables. In this way, the latent structure for the explanatory variables accounts for the country-specific heterogeneity and the dependence among covariates, while the concomitant variables makes the probability of belonging to a certain cluster depending on the initial conditions in the socio-economics structure.

Our findings provide evidence of strong presence of unobserved heterogeneity across countries. In fact, once the prior probabilities are conditioned on the initial level of the real GDP per capita, the tax burden and the education quantity, countries are clustered in six groups in which the homogeneity assumption about corruption determinants holds. The Romano test and Penalized likelihood criteria confirm the presence of heterogeneity. Moreover, the empirical distribution function shows that the data are satisfactorily fitted by the finite mixture model. Our findings show that the effect of the role of the State, the most debatable in literature, strongly vary across subgroups. In particular, it is found to increase corruption perception in countries as Italy or Spain, and to decrease it in countries as Norway, UK, and Canada. We also provided evidence that the highest the initial level of GDP per capita, the highest the probability for each country of belonging to the virtuous group.

References

[1] Acemoglu, D., and Verdier, T., (2000). The choice between market failures and corruption, *American economic review*, 194–211.

[2] Ades, A., and Di Tella, R. (1999). Rents, competition, and corruption, *American economic review*, 89(4), 982–993.

[3] Adsera A., Boix, C., Payne, M., (2003) Are you being served? Political accountability and quality of government, *Journal of Law, Economics, and Organization*, 19 (2), 445–490

[4] Aitkin, M.A.(1997). Contribution to the discussion paper by S. Richardson and P.J. Green, *Journal of The Royal Statistical Society*, **B**, 59, 764–768.

[5] Aitkin, M.A.(1999). A general maximum likelihood analysis of variance components in generalized linear models, *Biometrics*, 55, 117–128.

[6] Aitkin, M.A., Francis, B., and Hinde J., (2005). *Statistical Modelling in GLIM*, *Oxford University Press, Oxford, 2nd edition.*

[7] Aitkin, M.A., and Rocci, R. (2002). A general maximum likelihood analysis of measurement error in generalized linear models, *Statistics and Computing*, 12(2), 163–174.

[8] Alfó, M. and Trovato, G. (2004). Semiparametric mixture models for multivariate count data, with application, *Econometrics Journal*, 7, 1–29.

[9] Alfó, M. and Trovato, G., and Waldmann R.J., (2008). Testing for country heterogeneity in growth models using a finite mixture approach, *Journal of applied Econometrics* 23 (4), 487–514.

[10] Ali, A. M., and Isse, H. S., (2003). Determinants of economic corruption: a cross-country comparison, *Cato Journal*, 22, 449–466.

[11] Anderson, J. E., and Marcouiller, D. (2002), Insecurity and the pattern of trade: an empirical investigation, *Review of Economics and statistics*, 84(2), 342–352.

[12] Anderson, T.W. and Darling, D.A. (1954). A Test of Goodness-of-Fit. *Journal of the American Statistical Association*, 49, 765-?769.

[13] Bliss, C., and Di Tella, R. (1997). Does competition kill corruption?, *Journal of Political Economy*, 105(5), 1001–1023.

[14] Braun, M., and Di Tella, R., (2004). Inflation, inflation variability, and corruption, *Economics and Politics*, 16(1), 77–100.

[15] Cuervo-Cazurra A., (2006). Who cares about corruption?, *Journal of International Business Studies*, 37 (6), 807–822.

[16] Dayton, C. M., and Macready, G. B. (1988). Concomitant-variable latent-class models, *Journal of the American Statistical Association*, 83(401), 173–178.

[17] DiCiccio, T. J., and Romano, J. P. (1988). A review of bootstrap confidence intervals, *Journal of the Royal Statistical Society, Series B*50 No. 3 338–354.

[18] Durlauf, S. N., P. A. Johnson and J. R. W. Temple (2005). Growth Econometrics, *Handbook of Economic Growth* , 1A, S.N. Durlauf and P. Aghion (Eds.), North-Holland.

[19] Fréchette, G. R., (2006). Panel data analysis of the time-varying determinants of corruption, *CIRANO Working Papers 2006s-28.*

[20] Friedman, E., Johnson, S., Kaufmann, D., and Zoido-Lobaton, P., (2000). Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries, *Journal of Public Economics*, 76, 459–93.

[21] Fisman, R., and Gatti, R. (2002). Decentralization and corruption: evidence across countries, *Journal of Public Economics*, 83(3), 325–345.

[22] Fisman, R., and Miguel E. (2007) Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets, *Journal of Political economy*, 115(6), 1020–1048.

[23] Gerring J. and Thacker S.C.(2004). political institutions and corruption: the role of unitarism and parliamentarism, *British Journal of Political Science* , 34, No. 2, 295–990.

[24] Goel, R. K., and Nelson, M. A. (1998). Corruption and government size: a disaggregated analysis, *Public Choice*, 97 (1-2), 107–120.

[25] Grünn, B., and Leisch, F. (2008). FlexMix version 2: finite mixtures with concomitant variables and varying and constant parameters, *Journal of Statistical Software*, 28 (4), 1–35.

[26] Gundlach, E., and Paldam, M., (2009). The transition of corruption: From poverty to honesty, *Economics Letters, Elsevier*, 103, No. 3, 146–148.

[27] Habib, M., and Zurawicki, L. (2002). Corruption and foreign direct investment, *Journal of international business studies*, 291–307.

[28] Husted B.W. and Instituto Tecnologico y de Estudios (1999). Wealth, culture and corruption, *Journal of International Business Studies* , 30 (2), 339–359.

[29] Kiefer, J. and Wolfovitz, J.,(1956). Consistency of the maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association*,73, 805–811.

[30] Kotera, G., Keisuke, O., and Sovannroeun S., (2012). Government size, democracy, and corruption: An empirical investigation, *Economic Modelling*, 29 (6), 2340–2348.

[31] Laird, Nan. (1978). Nonparametric maximum likelihood estimation of a mixing distribution, *Journal of the American Statistical Association* 73 (364), 805–811.

[32] Lambsdorff, J., (1999). Corruption in empirical research: a review, *Transparency International*, processed 6.

[33] Lambsdorff, J., (2006). Measuring corruption?the validity and precision of subjective indicators (CPI), *Measuring corruption* , 81–99.

[34] Lancaster, T. and Montinola, G. (2001). Comparative Political Corruption: Issues of Operationalization and Measurement, *Studies in Comparative International Development.*

[35] Lindsay, B.G. (1983a). The geometry of mixture likelihoods: a general theory, *Annals of Statistics*, 11, 86–94.

[36] Mauro M. P., (1995). Corruption and growth, *The Quarterly Journal of Economics*, 110 (3), 681–712

[37] McLachlan, G. J. (1987). On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture, *Applied Statistics*, 318–324.

[38] McLachlan G.J., and Peel D., (2000). Finite Mixture Models. *John Wiley & Sons, New York*.

[39] Mocan, N. (2008). What determines corruption? International evidence from microdata, *Economic Inquiry*, 46(4), 493–510.

[40] Paldam M. (2002). The cross-country pattern of corruption: economics, culture and the seesaw dynamics, *Eureopean Journal of Political Economics*, 18, 215–240.

[41] Robertson C. J., and Watson A., (2004). Corruption and change: the impact of foreign direct investment, *Strategic Mnagement Journal*, 25, 385–396.

[42] Romano, J. (1988). "A Bootstrap Revival of Some Nonparametric Distance Tests." *Journal of the American Statistical Association*, 83, 698–708.

[43] Rose-Ackerman, S., (1997). The Political Economy of Corruption : Causes and Consequences, *World Bank*, Washington, DC. World Bank. https://openknowledge.worldbank.org/handle/10986/11629 License: CC BY 3.0 Unported.

[44] Serra, D. (2004). Empirical Determinants of Corruption: A sensitivity analysis, *Global Poverty Research Group. Working Paper GPRG-WPS–012* .

[45] Shapiro, S. S., and Wilk M.B., (1964). An analysis of variance test for normality(complete samples), *Biometrika*, 52(3 and 4),591–611.

[46] Svensson J. (2005). Eight questions about corruption, *Journal of Economic Perspectives*, 19 (3), 19–42.

[47] Treisman D. (2000). The causes of corruption: a cross-national study, *Journal of Public Economics*, 76, 399-457.

[48] Wei, S.J., (1997). Why is corruption so much more taxing than tax? Arbitrariness kills, *National Bureau of Economic Research*, w6255.

Table 1: The impact of economic fundamentals on corruption

Economic variables	High Corruption	Low corruption
Income per capita	Braun and Di Tella (2004) Frechette (2006)	Paldam (2000) Husted (1999) Svensonn (2005)
Government size	Ali and Isse (2003) Goel and Nelson (1998)	Fisman and Gatti (2002) Adserá et al. (2003)
Government intervention	Treisman (2000) Acemoglou et al. (2000)	Friedman (2000)

Notes: *income per capita*: GDP per capita; *government size*: public expenditure in final goods as share of GDP; *government intervention*: estimated in the form of regulation, taxation; *competition*: openness to trade, proxy as share of imports in GDP.

Table 2: Mean of covariates classified for high, medium and low CPI

	High	Medium	Low
Risk of punishment			
Democracy	0.54	0.70	0.94
Opposition Vote Share	21.10	32.99	45.22
Status of civil and political rights			
Free	61	315	201
Partially Free	197	115	12
Not Free	93	52	0
Index average	2.09	1.45	1.06
Independent Judiciary	0.32	0.60	1.00
Government Expenditure (perc of GDP)	12.43	16.12	19.03
Economics Incentive			
Business Freedom	38.58	52.44	66.53
Financial Freedom	45.36	56.41	73.10
Openness to Trade	72.22	85.83	89.10
Fiscal Freedom	83.09	80.44	70.95
Socio Economics Factors			
Real GDP per capita	4276.38	12178.08	29788.26
Education (Male average schooling)	6.93	8.37	11.45
Geographical Classification			
Eastern Europe, post Sovietic Union	75	103	0
Latin America	78	86	8
Middle Est, North Africa	13	82	6
Sub-Saharan Africa	108	79	0
Western Europe, North America	1	58	181
East Asia	6	28	6
South Est Asia	36	27	12
South Asia	132	6	0
Caribbean	2	13	0

Notes: The dependent variable, the *corruption perception index*, is divided in three categories, high, medium and low according to the index score. An index between 0-3 is associated to high corruption, 4-6 with medium corruption, and 7-10 to low corruption. The table shows the mean value of the covariates affecting the corruption level, divided for the macro-area presented in the paper: *risk of punishment*, *economics incentive* and *socio economics factors*. The *geographical classification* shows the number of countries in each geographic area that are characterized by high, low or medium level of corruption.

Table 3: Corruption Perception Index: Summary Statistics

variable	**mean**	**sd**	**skewness**	**kurtosis**	**min**	**max**	**N**
cpi	4.6068	2.3255	.7927	2.3815	0.4	10	1046

Notes: The Table displays summary statistics for the corruption perception index, showing: mean, standard error, skewness, kurtosis, maximum and minimum value.

Table 4: GLS and OLS with FE result

Variable	Coefficient	GLS Stand. Error	Coeff	OLS with FE St. Error
trend	-0.0410	0.0136 **	-0.0003	0.0065
business	0.0217	0.0028 ***	0.0024	0.00124
finance	0.0177	0.0024 ***	0.0042	0.0015 **
status,Partfree	-0.5009	0.1115 ***	0.1395	0.1340
status,Notfree	-0.2504	0.1822	0.0912	0.1050
demo	-0.4799	0.1214 ***	-0.2905	0.1171 *
ovs	0.0051	0.0022 *	0.0015	0.0012
jud	0.8594	0.0975 ***	0.2727	0.0767 ***
gce	0.5719	0.1246 ***	0.1113	0.1323
openk	-0.0014	0.0780	0.1497	0.1314
educ_95	-0.5241	0.1394 ***		
rgdpch_95	1.1702	0.0588 ***		
fiscal_95	-0.0874	0.2743		
(Intercept)	-7.7376	1.4132 ***	3.2004	0.6517 ***
ℓ	-1692.219			
BIC	3481.602			
n	1046			

Significance level: $***$: 0.1% $**$: 1% $*$: 5% . : 10%

Notes: Dependent variable: Corruption Perception Index; *business*: business freedom indicator (Heritage Foundation); *finance*: financial freedom indicator (Heritage Foundation); *status*: categorical variable for the political and civil rights (Freedom House); *demo*: dummy variable on the Democracy (Cheibub, Gandhi and Vreeland database); *ovs*: total votal share of all the parties in the opposition (Database of Political Institution); *jud*: dummy variable on judiciary independence (Henisz Index); *gce*: government consumption expenditure as a share of GDP (World Bank Database); *openk*: openness to trade (Penn World Table); *educ_95*: average years of schooling for men aged 25 or over (University of Washington); *rgdpch_95*: real GDP per capita (Penn World Table); *fiscal_95*: fiscal freedom indicator (Heritage Foundation).
BIC$=-2\ell(.) + d\log(n)$, where ℓ is the log-likelihood, d the number of parameters and n the sample size.

Table 5: Penalized criteria for finite mixture model

	iter	converged	k	k0	logLK	AIC	BIC	ICL
2	23	TRUE	2	2	-1218.5318	2493.064	2631.740	2633.313
3	21	TRUE	3	3	-1009.9007	2107.801	2325.722	2329.087
4	24	TRUE	4	4	-881.3361	1882.672	2179.836	2194.763
5	35	TRUE	5	5	-812.2970	1776.594	2153.001	2165.143
6	29	TRUE	6	6	-742.9693	1669.939	2125.590	2141.418
7	21	TRUE	6	7	-794.5870	1773.174	2228.825	2248.232

Notes: K number of components; logLK, log-likelihood
AIC= $-2\ell(.) + d$
BIC= $-2\ell(.) + d\log(n)$
ICL= $BIC + entropy$
where d is the number of parameters and n is the sample size

Table 6: Finite Mixture Model: prior probabilities

K	**prior**	**size**	**post**	**ratio**
1	0.1571	158	280	0.564
2	0.0825	102	162	0.63
3	0.0871	115	163	0.706
4	0.0944	136	204	0.667
5	0.2228	195	305	0.639
6	0.356	340	455	0.747

Notes: K number of components; *prior*: probability to belonging to that group k; *size*: number of country belonging to that group k; *post*: number of country belonging to that group after estimation; *ratio*: the ratio between size and post.

Table 7: Finite Mixture estimation

Variable	1 Comp.	2 Comp.	3 Comp.	4 Comp.	5 Comp.	6 Comp.
(Intercept)	-0.8574 .	2.0479	2.8143	-2.3754	-1.3803 0***	1.0547 ***
trend	-0.0430 ***	0.1242 ***	0.0089	-0.0444 ***	-0.1178 ***	0.0012
economics incentive						
business	0.0152 ***	-0.0067	0.0142 ***	0.0102 **	0.0091 ***	0.0071 ***
finance	0.0073 **	0.0328 ***	0.0497 ***	-0.0049 .	0.0076 ***	0.0076 ***
risk of punishment						
status,Partfree	-0.6181 ***	-0.8302 *	-3.2011	-0.9347	-0.3111 ***	-0.5907 ***
status,Notfree	-1.4302 ***	0.2356	-2.4365	-1.499	-1.1663 ***	-0.1272
demo	-0.5762 9***	0.6422	-1.0685	-3.0838	-0.112	-0.0166
ovs	0.0115 ***	0.0119 *	-0.0021	0.0046	0.0097 ***	-0.0006
jud	0.4069 ***	0.5752 *	1.2455 2 **	3.9643 ***	0.6114 ***	0.3771 ***
gce	0.7118 ***	-2.2242 ***	0.0949 4	3.1077 8***	0.2067 .	0.4199 ***
openk	0.5895 ***	1.3709 ***	0.0663 9	0.1298 2	0.9787 ***	0.0226
σ_k	0.392	0.6797	0.5201	0.4091	0.3543	0.4689
ℓ	-742.9693					
n	1046					

24

Significance level: $***$: 0.1% $**$: 1% $*$: 5% . : 10%

Notes: Dependent variable: Corruption Perception Index; *business*: business freedom indicator (Heritage Foundation); *finance*: financial freedom indicator (Heritage Foundation); *status*: categorical variable for the political and civil rights (Freedom House); *demo*: dummy variable on the Democracy (Cheibub, Gandhi and Vreeland database); *ovs*: total votal share of all the parties in the opposition (Database of Political Institution); *jud*: dummy variable on judiciary independence (Henisz Index); *gce*: government consumption expenditure as a share of GDP (World Bank Database); *openk*: openness to trade (Penn World Table).

Table 8: Concomitant Effects

	2 Comp.	**3 Comp.**	**4 Comp.**	**5 Comp.**	**6 Comp.**
(Intercept)	-25.5949	-69.5700 *	-70.6190 *	-4.2521	7.5369
rgdpch_95	2.4270 *	4.9180 **	6.1325 **	-1.5626 *	-2.6403 ***
fiscal_95	0.9748	4.7713	1.9041	3.4193	1.5710
educ_95	-0.6948	0.6701	1.2017	1.5685	4.4596 ***

Significance level: *** : 0.1% ** : 1% * : 5% . : 10

Note: The dependent variable is the Corruption Perception Index. All the concomitant variables are taken fixed at their initial values. *educ_95*: average years of schooling for men aged 25 or over (University of Washington); *rgdpch_95*: real GDP per capita (Penn World Table); *fiscal_95*: fiscal freedom indicator (Heritage Foundation). The first component is the reference class.

Table 9: Countries' groups

Algeria	Belarus	Australia	Canada	Benin	Albania
Botswana	Belgium	Austria	Denmark	Cambodia	Argentina
Brazil	Colombia	Bahrain	Finland	Dominican Republic	Armenia
Burkina Faso	Italy	Chile	Germany	El Salvador	Azerbaijan
Cameroon	Malaysia	France	Ireland	Fiji	Bangladesh
Chad	Saudi Arabia	Israel	Netherlands	Ghana	Bolivia
Cyprus	Slovenia	Japan	Norway	Guinea	Bulgaria
Estonia	Spain	Kuwait	Oman	India	Burundi
Gabon	Tunisia	Portugal	Singapore	Jamaica	China
Greece	UAE	Switzerland	Sweden	Lithuania	Congo
Hungary	Uruguay	New Zealand	US	Madagascar	Croatia
Jordan			UK	Malawi	Czech Republic
Korea, South				Mauritania	Ecuador
Lebanon				Mexico	Egypt
Mali				Morocco	Gambia
Mauritius				Mozambique	Georgia
Niger				Namibia	Guyana
Peru				Poland	Honduras
South Africa				Senegal	Indonesia
Turkey				Sri Lanka	Iran
Nepal				Swaziland	Kenya
				Vietnam	Laos
				Yemen	Latvia
				Zimbabwe	Lesotho
					Moldova
					Mongolia
					Nicaragua
					Pakistan (1972-)
					Panama
					Paraguay
					Philippines
					Romania
					Russia
					Rwanda
					Slovakia
					Syria
					Tajikistan
					Tanzania
					Thailand
					Uganda
					Ukraine
					Uzbekistan
					Venezuela
					Zambia

Table 10: Summary Statistics

	1 Comp.		**2 Comp.**		**3 Comp.**		**4 Comp.**		**5 Comp.**		**6 Comp.**	
	Mean	**St.Err**	**Mean**	**St.Err**	**Mean**	**St.Err**	**Mean**	**St.Err**	**Mean**	**St.Err**	**Mean**	**St.Err**
cpi	4.178	1.181	5.073	1.288	7.520	1.215	8.642	0.854	3.383	0.855	2.769	0.715
fiscal	80.407	8.066	76.871	12.723	75.471	10.342	70.245	12.055	80.765	7.298	83.900	6.337
demo	0.595	0.492	0.637	0.483	0.930	0.256	0.882	0.323	0.636	0.482	0.638	0.481
status	1.532	0.711	1.725	0.834	1.070	0.256	1.147	0.431	1.662	0.702	1.879	0.721
jud	0.462	0.500	0.500	0.502	0.965	0.184	0.971	0.170	0.426	0.496	0.479	0.500
business	52.030	14.134	54.076	14.285	62.425	13.448	68.528	16.794	44.060	16.522	41.650	16.575
educ	7.684	2.900	8.763	1.879	10.459	1.944	11.417	1.918	6.323	2.614	8.269	2.667
openk	72.836	38.018	99.855	53.950	64.474	29.631	102.790	89.029	80.475	29.330	79.174	36.970
rgdpch	9933.697	6492.986	19124.695	9753.745	25861.541	6735.981	31152.730	5700.395	5410.012	4248.530	5675.590	4043.584
gce	15.056	4.855	17.314	3.824	18.184	4.625	19.785	4.641	13.002	4.802	13.891	4.491
finance	56.899	15.530	54.412	16.970	66.522	18.451	71.985	15.339	49.487	16.614	50.147	18.869
ovs	29.140	21.050	26.036	22.815	37.046	17.712	48.043	14.038	30.576	22.365	26.241	21.028
N	158		102		115		136		195		340	

Notes: *cpi* is the Corruption Perception Index.*business*: business freedom indicator (Heritage Foundation); *finance*: financial freedom indicator (Heritage Foundation); *status*: categorical variable for the political and civil rights (Freedom House); *demo*: dummy variable on the Democracy (Cheibub, Gandhi and Vreeland database); *ovs*: total votal share of all the parties in the opposition (Database of Political Institution); *jud*: dummy variable on judiciary independence (Henisz Index); *gce*: government consumption expenditure as a share of GDP (World Bank Database); *openk*: openness to trade (Penn World Table). *educ*: average years of schooling for men aged 25 or over (University of Washington); *rgdpch*: real GDP per capita (Penn World Table); *fiscal*: fiscal freedom indicator (Heritage Foundation). *N* is the number of observations in each component.

Table 11: Reported *p-value* for the bootstrapping likelihood ratio test

Hypothesis Test		**p-value**
$H_0 : k = 1$	$H_1 : k \geq 3$	0.0102
$H_0 : k = 2$	$H_1 : k \geq 3$	0.01818
$H_0 : k = 5$	$H_1 : k \geq 6$	0.02381
$H_0 : k = 6$	$H_1 : k \geq 7$	0.3793
$H_0 : k = 6$	$H_1 : k \leq 5$	0.9659
$H_0 : k = 7$	$H_1 : k \leq 6$	0.14

Figure 1: Q-Q Plot: Corruption Perception Index

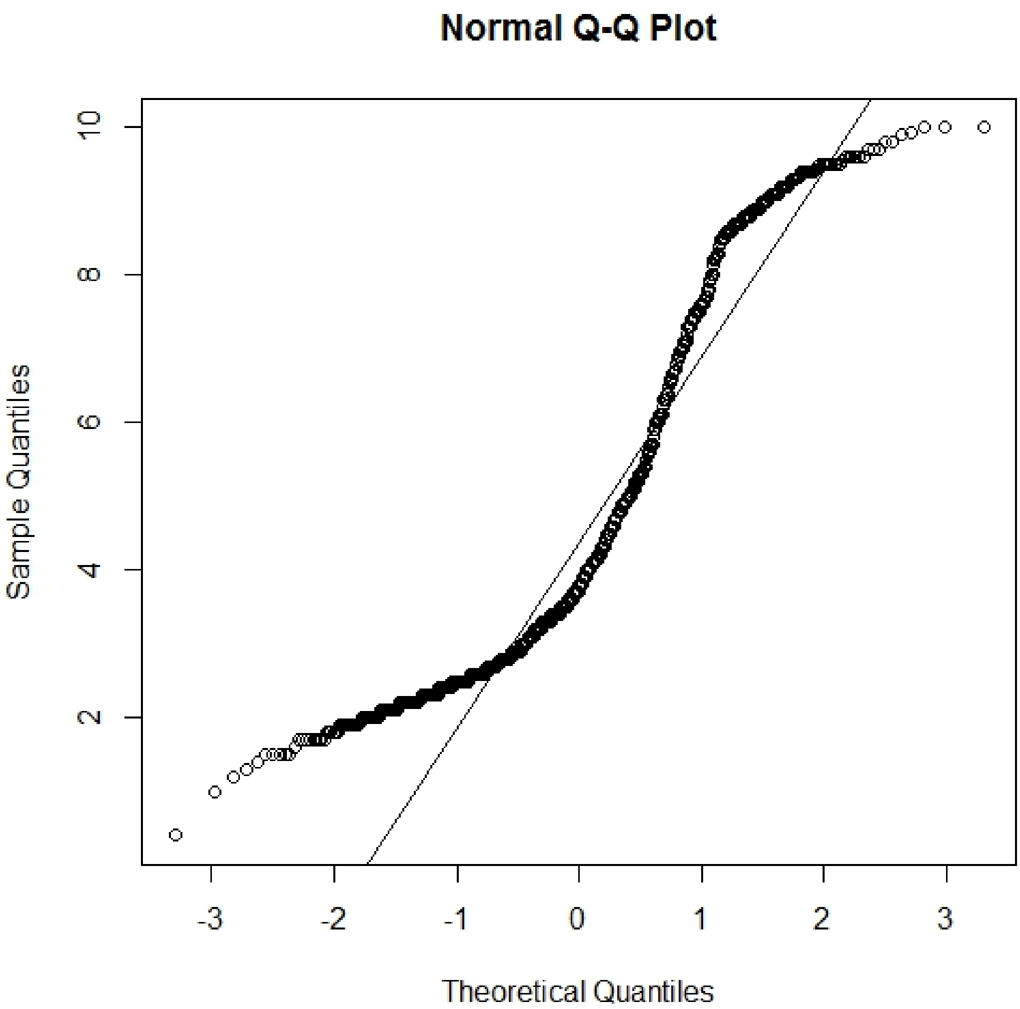


Figure 2: Density function: Corruption Perception Index

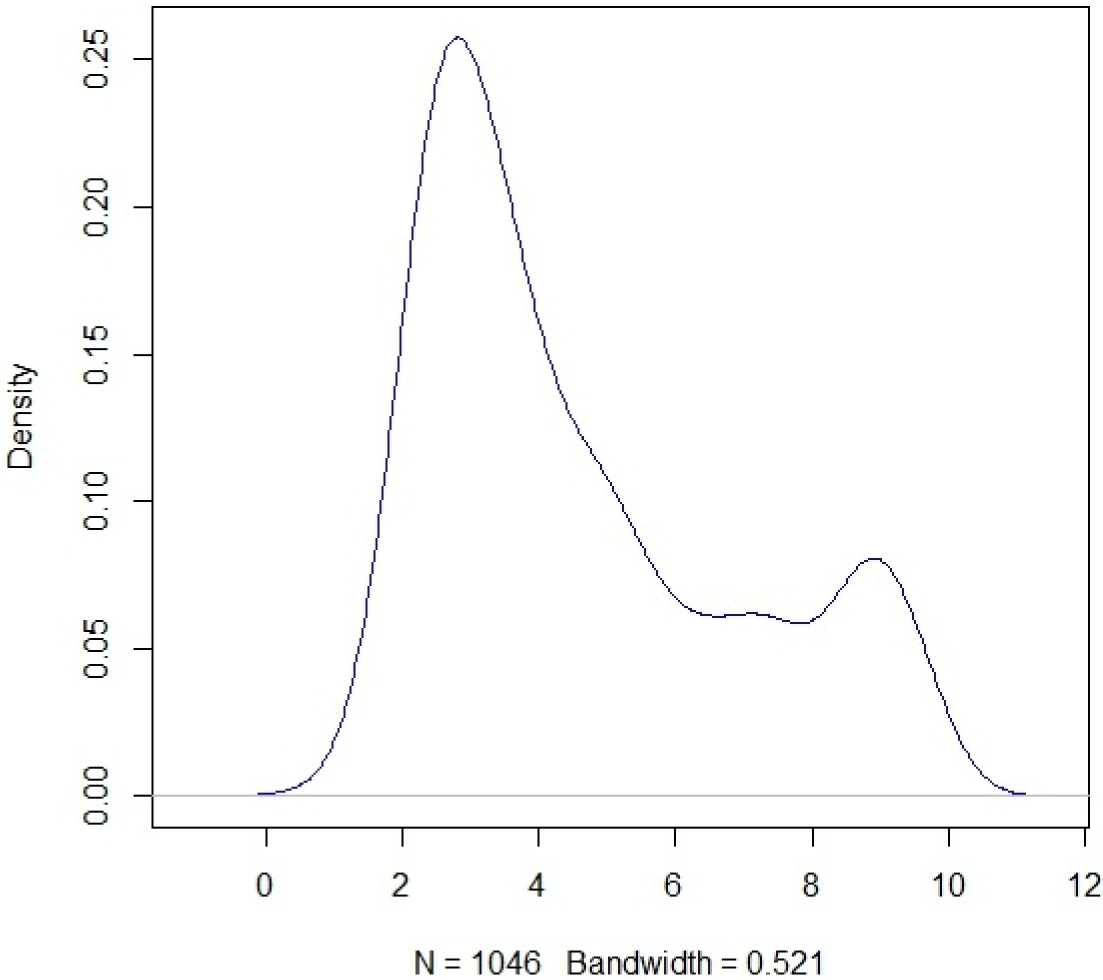


Figure 3: Kernel Distribution

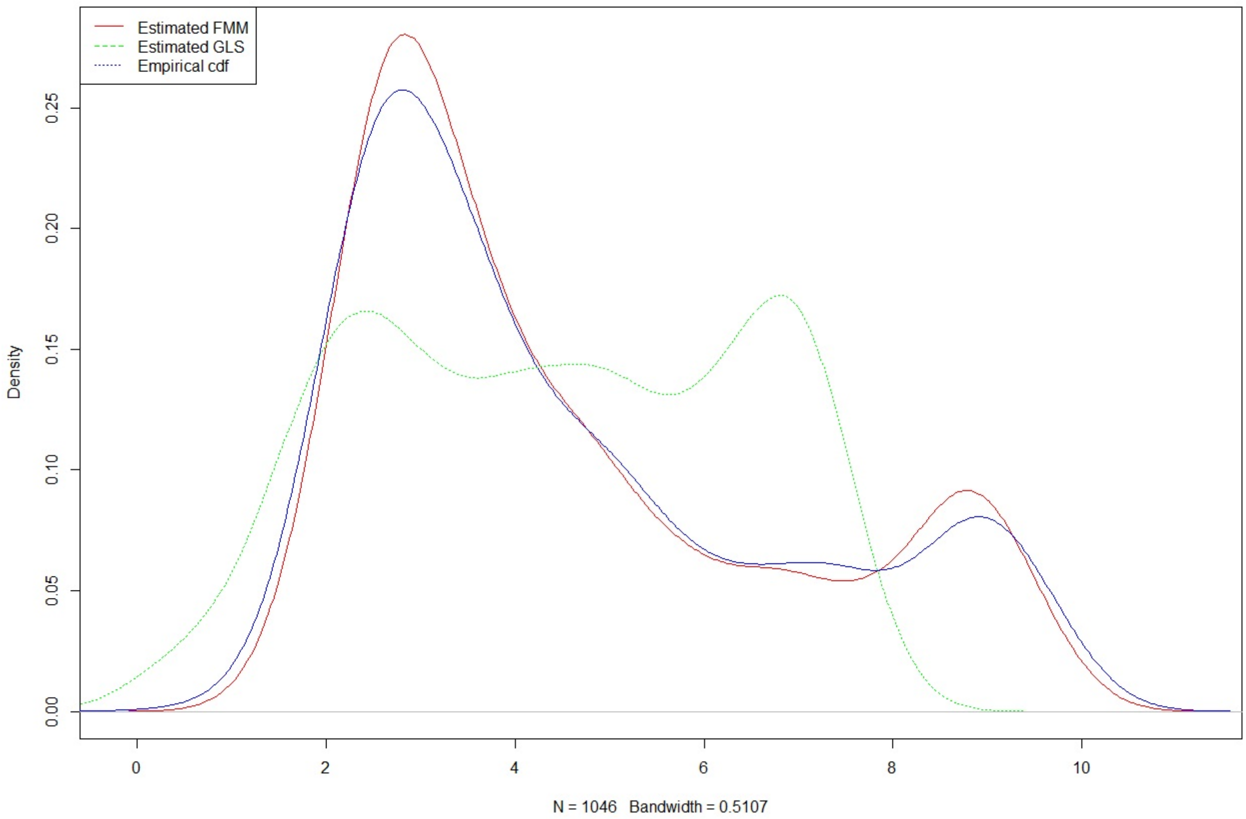


Figure 4: Components number: AIC BIC ICL criteria

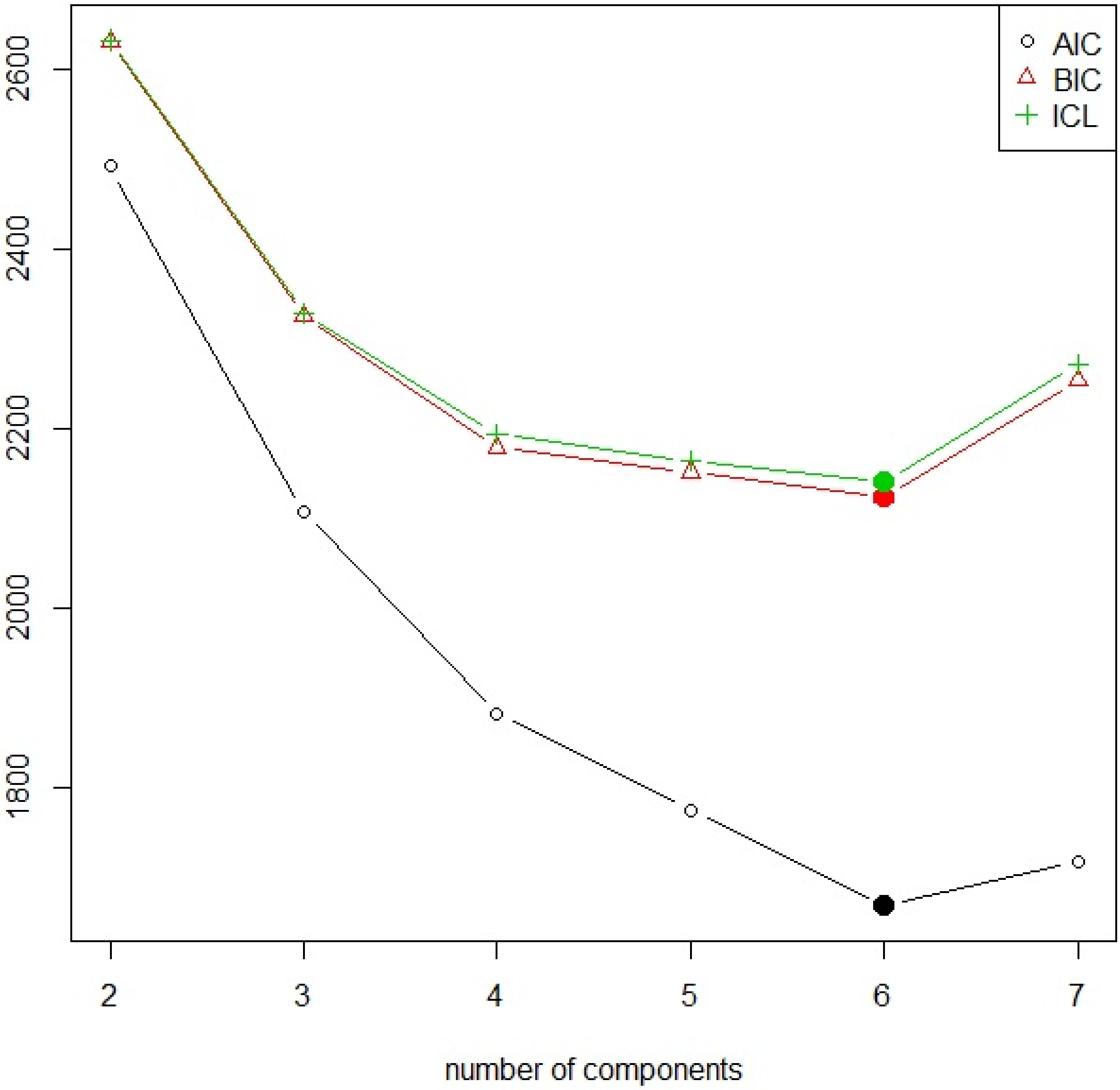
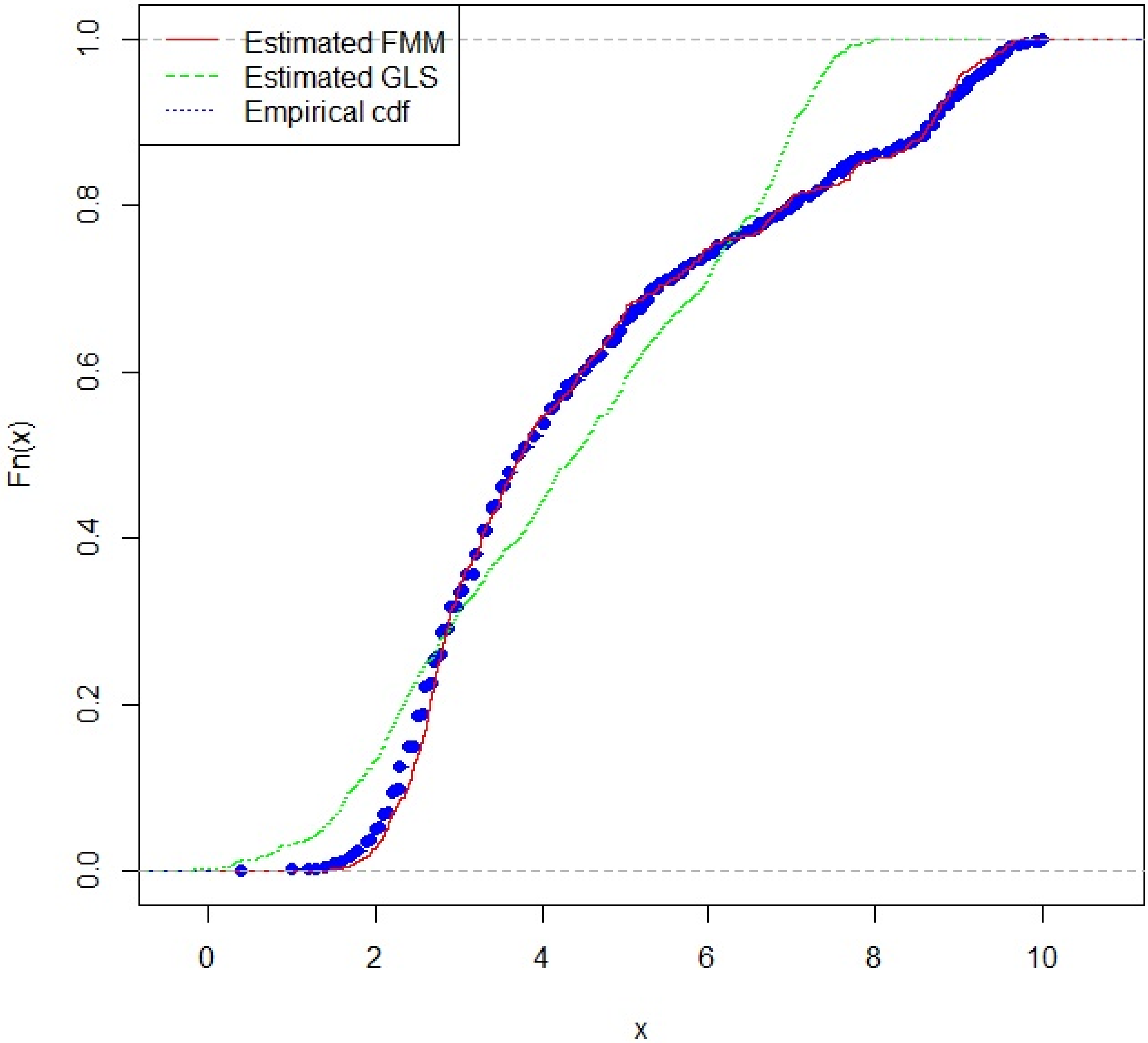


Figure 5: Estimated Cumulative Distribution Function



Acknowledgments

I am grateful to my supervisor Giovanni Trovato. He supported me since the very beginning of my academic career, starting from the bachelor degree. I thank him for his neverending encouragements, promptness and time during these years. He has been and still is an excellent mentor and his advices and support were precious to grow up as a researcher.
Thanks to Antonello Maruotti for being a precious guide during my visiting period at the University of Southampton. His suggestions were determinant for the development of my researching skills.
I wish to thank my parents, my brother Simone and my sister Martina for their essential moral encouragements. Particularly, I thank my mother, without her the pathway would have been much harder.
I am grateful to my “family abroad”, Damien, Domenico and Giulia, for making my visiting programme a wonderful experience and for their affinity.
I thank my good friend and colleague Gloria for our (frequent!) coffee breaks in the Department and the “vegan aperitivi” outside. This time with her was precious for overcoming difficult moments.
Thanks to my DEDI and DEF colleagues, in particular to my “historical” officemate Paola and my “new” officemate Andrea, for alleviating the routine in the Department, sharing together our everyday lives.
I would like to thank my “old” colleagues and above all friends, Sara, Salvatore and Claudio, and my everlasting friends Valerio and Antonella. Each of them, in their own special way, were an essential part of my University and everyday life from the very beginning.
Last but not least and/or *dulcis in fundo*, an enormous thank goes to Gionata. Happily and fortunately, he entered into my life during the last months. His being placid definitely turned the final stressful moments around, giving me a good dose of calm, relax, and joy.

Rome January 20, 2015