On the Determinants of Growth Volatility*

Davide Fiaschi - Andrea Mario Lavezzi
University of Pisa
Dipartimento di Scienze Economiche
Via Ridolfi 10, 56124 Pisa, Italy
E-mail: dfiaschi@ec.unipi.it
E-mail: lavezzi@ec.unipi.it

October 23, 2003

Abstract

We propose a model where the growth rate volatility of a country is explained by structural change and the size of the economy. We test these predictions by adopting two measures of growth volatility: the standard deviation of growth rates and some indices based on Markov transition matrices. Both methods lead to the same results: growth volatility appears to (i) decrease with total GDP, (ii) increase with the share of the agricultural sector on GDP. Trade openness can also play a role in conjunction with total GDP. In accordance with our model, the explanatory power of per capita GDP, a relevant variable in other empirical works, vanishes when we control for these variables.

Keywords: growth volatility, Markov transition matrix, structural change, nonparametric methods.

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

*We are very grateful to Carlo Bianchi, Eugene Cleur, Simon Woods and seminar participants at Siena and Guanajuato for helpful comments. The usual disclaimers apply.

1 Introduction

The relationship between income level and growth rate volatility (*GRV* henceforth) has received little attention up to now. Contributions can be divided into two main groups. The first highlights that development is accompanied by a sharp reduction in *GRV* (see Acemoglu and Zilibotti (1997) and Pritchett (2000)), while the second refers to a negative relationship between the *size* of an economy and *GRV* (see Canning *et al.* (1998)).[1]

Since development is generally intended as an increase in per capita GDP, a first possible empirical investigation regards the relationship between *GRV* and per capita GDP. In this light, we analyze structural change, a typical phenomenon associated to development. In fact, a plausible explanation of the reduction in *GRV* as development proceeds resides in the decreasing weight of sectors with more volatile output, like agriculture and primary sectors, with respect to sectors with less volatile output, like manufacturing and services.[2] Differently, the increase in the number of sectors (or productive units) associated to a growing size of the economy is the most common explanation of the relationship between the size of the economy and *GRV*. In fact, a reduction in aggregate *GRV* may derive from averaging an increasing number of sectoral growth rates, since idiosyncratic shocks to each sector would tend to cancel out by the law of large numbers.

We test for the existence of these relationships in a large sample of countries from Maddison (2001)'s dataset. In particular we focus on the effect of three variables on *GRV*: (i) the level of per capita GDP (*GDP* henceforth) as proxy of the level of development, (ii) the share of agriculture on GDP (*AS* henceforth) as proxy of structural change and (iii) total GDP (*TGDP* henceforth) as proxy of the size of the economy. We also consider a measure of trade openness (*TR* henceforth), to proxy the effective dimension of an economy which may not be entirely captured by *TGDP* only.

Individually, we find an inverse relationship between *GRV* and both *GDP* and *TGDP*, and a positive relationship between *GRV* and *AS* as we expected, although some nonlinearities appear in the latter case. *TR* shows a nonlinear behavior, but we argue that the effect of this variable on *GRV* has to be evaluated jointly with *TGDP*. When we consider all the variables, *TGDP* explains the largest part of growth volatility. In particular, the effect of *GDP* on *GRV* vanishes when it is considered jointly with *TGDP*, *TR* and *AS*. These findings agree with the predictions of our model, in which *GRV* is explained by structural change and, especially, by the extent of the economy.

[1] Acemoglu et al. (2003) highlight another possible causal explanation of volatility based on the lack of strong "institutions" (e.g. enforcement of property rights, corruption, political instability), while Easterly et al. (2000) focus on the development of the financial sectora as a cause for the reduction in volatility.

[2] So far the literature on structural change has not paid attention to this issue (see e.g. Pasinetti (1981)).

From the theoretical point of view, our work is also related to papers such as Scheinkman and Woodford (1994) and Horvath (1998), which study the emergence of aggregate fluctuations from local shocks. None of them is however explicitly concerned with structural change. Acemoglu and Zilibotti (1997) study an economy where an increasing number of sectors allows for a diversification of investment, and is associated to a reduction in aggregate *GRV*. A direct implication is that risk-adverse agents, by investing in more productive and more risky sectors, determine an increase in the growth rate.[3] Hence their approach differ from ours as we focus on the specificity of the sectors (agricultural sector vs other sectors) and not only on their number. Moreover, they do not explicitely interpret the number of sectors as a proxy of the size of the economy.

As a first step in our empirical analysis we follow the Canning *et al.* (1998)'s approach, where all observations are pooled and then partitioned in classes. We measure *GRV* for each class of *GDP*, *AS* and *TGDP* as the standard deviation of growth rates associated to the observations in each class. We estimate by nonparametric methods (this is a crucial difference with respect to Canning *et al.* (1998)) the relationship between *GRV* and our explanatory variables, exploring in particular the effects of their interactions. Here *GDP* appears to play no role in the explanation of *GRV* when *AS*, *TGDP* and *TR* are included in the regression.

However, we argue that a drawback of the procedure based on pooling is to ignore the relevant information on the dynamics of individual countries. Therefore we present a new statistical methodology based on Markov transition matrices. In particular we propose some *growth volatility indices* based the literature on mobility indices (see, e.g. Bartholomew (1982)). We reinterpret a set of indices generally utilized to measure intergenerational mobility as measures of volatility, and propose two new indices. By applying these indices to our sample, we find a confirmation of the previous findings. This methodology hardly allows for a rigorous statistical analysis of the joint effect of our variables on growth volatility, because of the limited availability of data, but we provide some intuition supporting the result that *GDP* is not informative, in presence of the other variables.

The paper is organized as follows. Section 2 proposes a simple model to explain the growth volatility of a multisector economy. Section 3 contains a nonparametric data analysis of *GRV*; Section 4 introduces the *GRV* indices; Section 5 presents and discusses the results of the calculation of these indices; Section 6 concludes.

[3]Here we are not interested in the link between *GRV* and long-run growth as in Ramey and Ramey (1995).

2 A Basic Analytical Framework

In this section we present a simple model to highlight the key factors which can account for *GRV* in a country. In particular our focus is on the composition of output and the size of the economy.

Consider an economy with N_t sectors, where t indexes time. Sector i's output grows according the following rule:

$$y_t^i = y_{t-1}^i \left(1 + g_t^i \varepsilon_t^i\right),$$

where y_t^i is output in period t of sector i, g_t^i is the exogenous growth rate of sector i, and ε_i^t is a random shock.

We assume that random shocks are normally distributed with mean 1 and variance $\left(\sigma^i\right)^2$, that is:

$$\varepsilon_t^i \sim N\left(1, \left(\sigma^i\right)^2\right).$$

Let Γ_t be the $N_t \times N_t$ covariance matrix, where γ_t^{ij} is an element. Notice that assuming a nonzero covariance among shocks is a simple way to model sectoral interdependence.[4] We assume that the autocorrelation of the shocks is zero, that is $cov\left(\varepsilon_t^i, \varepsilon_{t-1}^i\right) = 0$, $\forall i = 1, ..., N_t$ and $\forall t$. Finally, we assume that $\sigma^{i-1} \geq \sigma^i$, $i = 2, ..., N_t$, that is we order sectors on the basis of *GRV*.[5]

Notice that shocks are assumed to be normally distributed for analytical convenience. In fact, this allows us to measure aggregate *GRV* by the standard deviation of the aggregate growth rate. If we relax this assumption, measuring *GRV* of a country can become complex. We return on this point in the sections devoted to the empirical analysis.

Let Y_t be aggregate output in period t, that is:

$$Y_t = \Sigma_{i=1}^{N_t} y_t^i.$$

Therefore the aggregate growth rate is given by:

$$\mu_t = \frac{\Sigma_{i=1}^{N_t} y_{t-1}^i \left(1 + g_t^i \varepsilon_t^i\right)}{\Sigma_{i=1}^{N_t} y_{t-1}^i} - 1 = \Sigma_{i=1}^{N_t} \alpha_{t-1}^i g_t^i \varepsilon_t^i, \tag{1}$$

where $\alpha_{t-1}^i = \frac{y_{t-1}^i}{\Sigma_{i=1}^{N_t} y_{t-1}^i}$ is the share of output of sector i with respect to total output, so that $\Sigma_{i=1}^{N_t} \alpha_{t-1}^i = 1$, $\forall t$.

[4] Horvath (1998) shows that a multisector model with intermediate goods and idiosyncratic shocks to individual sectors can generate an aggregate dynamics where sectoral outputs are correlated.

[5] Grossman and Kim (1996) endogenize the different volatility of sectors on the basis of rent-seeking theory. Here, we argue that these sectors are intrinsically more subject to random shocks, e.g. changes in terms of trade, climatic changes and the like.

From definition (1) we have that the expected value and variance of μ_t are given by:

$$\bar{\mu}_t = E_t [\mu_t] = \Sigma_{i=1}^{N_t} \alpha_{t-1}^i g_t^i \tag{2}$$

$$\bar{\sigma}_t^2 = E_t \left[\left(\Sigma_{i=1}^{N_t} \alpha_{t-1}^i g_t^i \eta_t^i \right)^2 \right], \tag{3}$$

where $\eta_t^i = \varepsilon_t^i - 1$. Trivially, η has the same properties of ε, but its mean is equal to 0 (that is $\eta_t^i \sim N\left(0, \left(\sigma^i\right)^2\right)$). It follows that μ_t is normally distributed, that is $\mu_t \sim N\left(\bar{\mu}_t, \bar{\sigma}_t^2\right)$. From (3) we obtain the following expression for $\bar{\sigma}_t^2$:

$$\bar{\sigma}_t^2 = \Sigma_{i=1}^{N_t} \left(\alpha_{t-1}^i g_t^i \sigma^i \right)^2 + \Sigma_{i=1}^{N_t} \Sigma_{j=1, j \neq i}^{N_t} \alpha_{t-1}^i \alpha_{t-1}^j g_t^i g_t^j \gamma_t^{ij}, \tag{4}$$

where $\gamma_t^{ij} - 1$ is the covariance between η_t^i and η_t^j.

The functional form of Equation (4) does not allow for a simple identification of the effects of the elements on the right-hand side on $\bar{\sigma}_t^2$, except for g_t^i. An increase of g_t^i, *ceteris paribus*, increases both $\bar{\mu}_t$ and $\bar{\sigma}_t^2$, i.e. $\frac{\partial \bar{\sigma}_t^2}{\partial g_t^i} > 0$, $\forall i$. However, the effects of the other variables involved, in particular the number of sectors N_t and of structure of economy $\left(\alpha_{t-1}^1, ..., \alpha_{t-1}^{N_t} \right)$, may not be so easily identifiable.

To proceed, suppose that Y_t comes from the agricultural sector A (sector 1), and from the rest of economy R (sectors $2, ..., N_t$), which includes secondary and tertiary sectors (we will use this distinction in the empirical analysis). Equation (4) becomes:

$$\bar{\sigma}_t^2 = \left(\alpha_{t-1}^A g_t^A \sigma^A \right)^2 + \left(\alpha_{t-1}^R g_t^R \sigma_t^R \right)^2 + \alpha_{t-1}^A \alpha_{t-1}^R g_t^A g_t^R \gamma_t^{AR}. \tag{5}$$

It is plausible to assume that $\gamma_t^{AR} = 0$ because shocks to A and R are likely to be of different nature and uncorrelated.[6] Therefore we have:

$$\bar{\sigma}_t^2 = \left(\alpha_{t-1}^A g_t^A \sigma^A \right)^2 + \left[\alpha_{t-1}^R g_t^R \sigma_t^R \right]^2. \tag{6}$$

Generally, a change in α_{t-1}^A and $\alpha_{t-1}^R = 1 - \alpha_{t-1}^A$, and/or a change in the number of sectors N_t have an ambiguous effect on aggregate variance. Let us analyze first the role of N_t.

Number of Sectors and Growth Volatility Some authors argue that the size of an economy, in terms of number of sectors or units of production, may affect aggregate *GRV* (e.g. Scheinkman and Woodford (1994)). In our model, the possible negative correlation between *GRV* and N can derive from an inverse correlation between σ^R and N. We can identify simple conditions under which $d\sigma^R/dN < 0$. Assume that $g_t^i = g^R$, $\gamma_t^{ij} = 0$,

[6] For a discussion of the relationship between $\bar{\sigma}^2$ and Γ see Horvath (1998).

$\alpha_0^i = \frac{1}{N_0 - 1}$, for $i, j = 2, ..., N_t$ and $\forall t$. Then, from Equation (4) written only for R, we have:

$$\left(\bar{\sigma}_t^R\right)^2 = \left(\frac{g^R}{N_t - 1}\right)^2 \left[\Sigma_{i=2}^{N_t} \left(\sigma^i\right)^2\right],$$

that is $\left(\bar{\sigma}_t^R\right)^2$ is decreasing in N_t and increasing in g^R, given that $\sigma^{i-1} \geq \sigma^i$.

Hence, the higher is the number of sectors in R, the lower is the variance of its growth rate, if the covariance between sectors is negligible (this is an application of the law of large numbers). If the size of an economy is positively related to the number of sectors N, then the size of an economy and its growth volatility are inversely related. Moreover, higher g^R leads to higher GRV but, if the output of some sectors has a strong positive correlation with the output of others, then GRV can nonetheless increase if the latter effect is stronger than the effect of the increase in N.[7]

To conclude, if $\sigma^R = \sigma^R(N)$, where $d\sigma^R/dN < 0$, then from Equation (6) we obtain:

$$\frac{\partial \bar{\sigma}_t^2}{\partial N_t} = 2\left[\alpha_{t-1}^R g_t^R\right]^2 \sigma^R \frac{d\left(\sigma^R\right)^2}{dN_t} < 0. \qquad (7)$$

We show below that this relationship finds an empirical support when we proxy for N by the dimension of the economy.

Composition of Output and Growth Volatility In a typical process of growth and structural change, primary sectors grow less than industrial and service sectors. This implies that the share of sectors with higher variance declines over time. The overall result would be a decrease in aggregate GRV, as the latter is a weighted sum of sectors' variances, and weights are proportional to sectors' shares.

From Equations (1) and (6) we have:

$$\bar{\sigma}_t^2 = \left(\alpha_{t-1}^A g_t^A \sigma^A\right)^2 + \left[\left(\mu_t - \alpha_{t-1}^A g_t^A\right)\sigma^R\right]^2.$$

Calculations lead to:

$$\frac{\partial \bar{\sigma}_t^2}{\partial \alpha_{t-1}^A} > 0 \Leftrightarrow \alpha_{t-1}^A > \frac{\mu_t}{g_t^A}\left[1 + \frac{\left(\sigma^A\right)^2}{\left(\sigma^R\right)^2}\right]^{-1} = \bar{\alpha}. \qquad (8)$$

This means that for $\alpha_{t-1}^A < \bar{\alpha}$ ($\alpha_{t-1}^A > \bar{\alpha}$) GRV is decreasing (increasing) in the share of the agricultural sector α^A. That is, the relationship between

[7]An example can be the emergence of a financial sector, whose output is correlated to many sectors through the capital market. This remark could introduce the very interesting question whether GRV remains stable over time given the same level of GDP. For example, the development of a global capital market may increase the interdependence among sectors and possibly GRV, without implying an increase in the level of GDP.

α^A_{t-1} and $\bar{\sigma}^2_t$ is U-shaped. Moreover, if $\sigma^R = \sigma^R(N)$ and $d\sigma^R/dN < 0$, then the threshold value $\bar{\alpha}$ decreases in N_t.[8]

To summarize our results consider the following equation, derived from Equation (6):

$$\bar{\sigma}^2_t = \left(\mu_t \sigma^R\right)^2 + \left(\alpha^A_{t-1} g^A_t\right)^2 \left[\left(\sigma^A\right)^2 + \left(\sigma^R\right)^2 - \frac{2\sigma^R}{\alpha^A_{t-1} g^A_t}\right]. \qquad (9)$$

In Equation (9) aggregate variance depends on two terms: the first term captures the effect of the variance of the "rest of the economy", which we argue depends negatively on the number of sectors N (see Equation (7)); the second term represents the effect of the share of agriculture α^A, whose sign depends in a non-trivial way on the interaction with N, via σ^R (see condition (8)). Notice finally that *GDP* does not play any role in the model. In our empirical analysis we estimate Equation (9).

3 Nonparametric Estimation

We use data on *GDP* and *TGDP* from Maddison (2001)'s database and data on agriculture and trade from the World Bank's *World Development Indicators 2002*. Our sample includes 119 countries for the period 1960–1998.[9] As noted, we proxy for the structure of the economy by the share of the agricultural sector in aggregate value added, *AS*, and measure the *effective* dimension of the economy, related to the number of sectors N in the model, both by the *total* GDP (*TGDP*) *and* trade openness (*TR*), which is the ratio of the sum of imports and exports on GDP. The latter, jointly with *TGDP*, would provide a more exact measure of the extent of the overall market for an economy.

We consider both the cross-country and the time-series dimension of growth volatility. In particular, to evaluate the relation between *GRV* and level of development we separate all observations on *GDP* and *TGDP* into 151 classes with a similar number of observations (approximately 30), while to evaluate the relation between *GRV* and structural change we separate all observations on *AS* into 109 classes. Finally, for the relation between *GRV* and *TR* we have 125 classes of observations.

For every observation on *GDP* in year t we calculate the growth rate

[8] Notice that the U-shaped relation between α^A_{t-1} and $\bar{\sigma}^2_t$ resembles the relation between the variance of a portfolio and the share of the more volatile asset. In the problem of portfolio choice, the variance of portfolio decreases with the share of the more volatile asset until a positive threshold value is reached, then increases.

[9] Data on *GDP* and *TGDP* are in 1990 international dollars. Not all observations on agriculture and trade openness were available for each country for all years. See Appendix A for the country list.

from t to $t+1$.[10] Figure 1 reports the standard deviation of growth rates, *STD*, relative to the observations in a class against, respectively, the log of the average *GDP*, *AS*, *TGDP* and *TR* in that class, and run a nonparametric estimation of these relationships.[11]

Figure 1 is the counterpart of Figure 1 in Acemoglu and Zilibotti (1997), where only cross-country variation in growth volatility is considered. They estimate an OLS regression and find a decreasing relationship betweeen growth volatility and development, proxied by the initial level of *GDP*.

In our case, we see at first glance that *GRV* tends to fall with *GDP*. The high volatility at the lowest and, especially, highest *GDP* levels is associated with a much wider variability band, meaning that there the estimate is not precise. In Figure 1 growth volatility appears to be increasing with *AS*. This relation is not monotonic, but the variability band is tighter where the upward sloping portion is steeper, indicating that the estimation is more precise where the curve is sharply increasing (we return on this below). In Figure 1 *GRV* clearly decreases with *TGDP*, as the extreme portions of the estimate have a wide variability band.

Finally, the relationship between *GRV* and *TR* in Figure 1 appears inversely U-shaped. In particular, the estimate of both the decreasing parts has a wide variability band. As noted, the impact of *TR* on *GRV* does not interest us *per se*, but in conjunction with *TGDP* when we proxy for the effective size of an economy. In our view, the effective size of the economy increases if it is highly integrated with other economies.

Notice that in Figure 1 we have studied the effects on *GRV* of the variables taken individually. However, from our model, these variables are expected to have a joint effect on *GRV*, that is their effect should be evaluated given the presence of other variables and of possible interactions among them.

To test the implications of Equation (9) we estimate the following gen-

[10] For data on *AS*, *TGDP* and *TR* we consider the corresponding observation on *GDP* and calculate the associated growth rate.

[11] The nonparametric estimate is obtained with the statistical package included in Bowman and Azzalini (1997). We used the standard settings suggested by the authors (i.e. optimal normal bandwidth). To test the robustness of this estimate, we ran an alternative nonparametric regression using the plug-in method to calculate the kernel bandwidth, and obtained a similar picture. We refer to Bowman and Azzalini (1997) for more details. We report the variability bands representing two standard errors above and below the estimate. They give a measure of the statistical significance of the estimate (see Bowman and Azzalini (1997), pp. 29–30 for details on variability bands vs confidence bands). Data sets and codes used in the empirical analysis are available on the authors' websites (http://www-dse.ec.unipi.it/fiaschi and http://www-dse.ec.unipi.it/lavezzi).

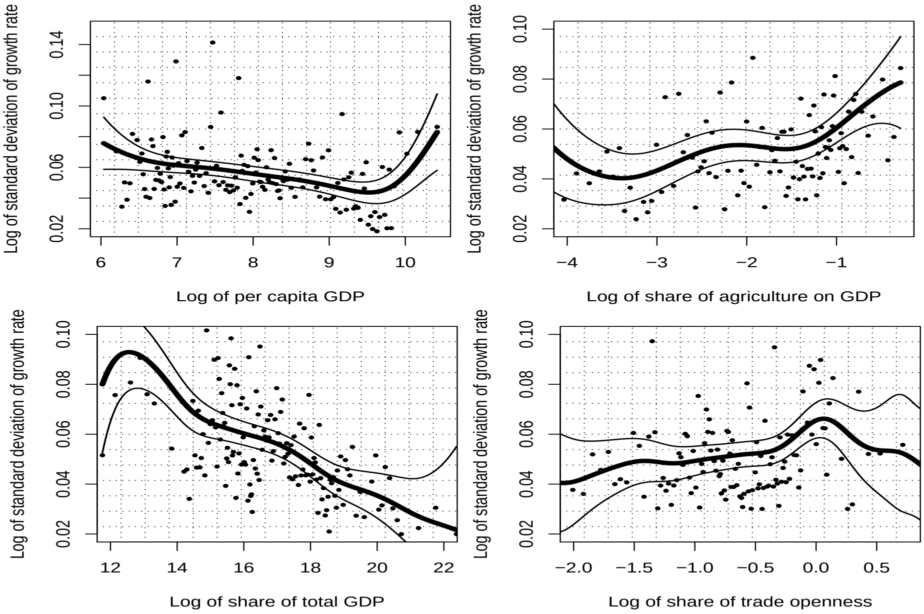

Figure 1: GRV estimated by STD vs, respectively, log of GDP, AS, $TGDP$, and TR

eralized additive models:[12]

$$STD_i = \beta_0 + \sum_{j \in P_1} s_j(x_{ji}) + \sum_{q \in P_2} \sum_{j \in P_3, j \neq q} s_{j,q}(x_{ji}, x_{qi}) + \quad (10)$$

$$+ \sum_{k \in P_4} \sum_{q \in P_5} \sum_{j \in P_6, j \neq q \neq k} s_{k,j,q}(x_{ki}, x_{ji}, x_{qi}) + \epsilon_i \quad (11)$$

where STD_i indicates that standard deviation in class i, $P_z \subseteq \{GDP, AS, TGDP, TR\}$, and $s_j(.)$, $s_{j,q}(.)$ and $s_{k,j,q}(.)$ are functions to be estimated nonparametrically. Functions $s_{j,q}(.)$ and $s_{k,j,q}(.)$ capture the effect of the interactions among the explanatory variables x_{ki}, x_{ji} and x_{qi}. Here GDP is considered to check the robustness of the results to its inclusion, and for comparison with existing results in the literature (remember that in our theoretical model GDP does not affect GRV).

Estimation by generalized additive model is particularly well- suited in this contest because it is not affected by multicollinearity, a potential problem given the high correlation between GDP, AS, $TGDP$ and TR.[13]

We estimated Equation (10) alternatively with classes defined on the

[12] As we discussed above, in this paper we focus only on structural change and the size of economy. Hence we do not consider in the empirical analysis the covariance matrix Γ and growth rates of individual sectors.

[13] For example, the coefficient of correlation between $TGDP$ and AS is -0.79, while for $TGDP$ and TR it is equal to -0.69.

basis of *GDP* and *TGDP*. The best results obtains with *TGDP*, and are reported in Table 1.[14]

For every estimated model we report the p-value for the approximate significance of each individual explanatory variable, the estimated degrees of freedom, the *GCV* score, and the value of R^2. Model 1 in Table 1 directly corresponds to Equation (9). From this equation we expected an effect on *GCV* from the number of sectors alone and from an interaction between N and *AS*. We consider the interaction between *TGDP* and *TR* as a proxy for N. These effects are indeed highly significant, and the specification of Model 1 produces the best results, in particular for the *GCV* score. The interaction between *TGDP* and *TR* and the interaction between these variables and *AS* account for 65% of the variance of *STD*.

Models 2−6 test the robustness of this result to alternative specifications. In Model 2 we check whether the inclusion *TR* affects the results. We see that the *GCV* score increases while R^2 decreases, although the two variables are highly significant. Therefore we conclude that *TR* should be included. In Model 3 we check for the relevance of *AS*, as its effect may be completely captured by the size of the economy. For instance, it is likely that an economy with a large agricultural share is quite underdeveloped and has a small size. However, with respect to the results of Model 1 we see that the exclusion of *AS* worsen the results.

Given that *TR* and *AS* are relevant, we check in Model 4 for the exclusion of their interactions. We can see that the only significant variable is *TGDP*, and that the results are worse than in Model 1. Hence, we conclude that the importance of *TR* and *AS* lies in their interactions. In Model 5 we check whether these variables are relevant when taken in one single interaction term, and conclude in the negative. Finally, in Model 6 we add *GDP* to our best specification, Model 1. We see that *GDP* is not significant, while the significance of the other terms is preserved.

Therefore, we argue that the effect of "development", when measured by *GDP*, on the decrease in *GRV* is broadly ascribable to our variables proxing for the dimension of the economy and structural change. Hence, it seems that other potentially relevant factors whose effect might be captured by *GDP* (e.g. the development of a financial system or of other "stabilizing" institutions), are not actually informative in presence of our variables.[15]

[14]Results obtained when classes are defined using *GDP* are available upon request. The smooth terms $s(.)$ in Equation (10) are represented by penalized regression splines. The smoothing parameters are chosen to minimize the *Generalized Cross Validation* score (*GCV*) of the model, and the estimated degrees of freedom are computed as part of the minimization process (see Wood (2000) for details).

[15]In Appendices B, C and D we show that these results are largely robust to different definitions of *STD*. The main differences are: i) the relevance of *TR* seem to depend on the definition of *STD*; ii) *GDP* is significant with cross-section data. However, in the latter case the number of available data is quite low. In particular, data on many developing countries are missing and therefore the estimation misses important phases of development

Model	1	2	3	4	5	6
Constant	0	0	0	0	0	0
$s(TR, TGDP)$	0 (30.37)	-	0 (14.18)	-	-	0 (30.45)
$s(AS, TGDP)$	-	0.02 (19.99)	-	-	-	-
$s(AS, TR, TGDP)$	0 (7)	-	-	-	0 (22.15)	0 (7)
$s(TGDP)$	-	0 (8.52)	-	0 (4.10)	-	-
$s(AS)$	-	-	-	0.32 (1)	-	-
$s(TR)$	-	-	-	0.73 (1.02)	-	-
$s(GDP)$	-	-	-	-	-	0.42 (2.28)
GCV score(*10^{-4})	3.3733	3.5599	3.6217	3.6204	3.8237	3.4117
R^2	0.65	0.57	0.45	0.38	0.49	0.66
Number of obs.	149	149	149	149	149	149

Table 1: Estimation of Equation (10). Dependent variable is STD, classes defined in terms of TGDP. The p-value of the terms and the estimated degrees of freedom (in parenthesis) are reported

Figure 2 reports the estimated effects of the individual variables on STD, on the basis of Model 1.[16]

Estimation highlights that STD has a significant positive correlation with AS; moreover, a clear negative relationship exists between STD and $TGDP$ that, except for the lowest values of $TGDP$ where the number of observations is low. Finally, the effect of TR on STD appears to be relevant only for low values of TR, but its sign is ambiguous.

This approach has the drawback of ignoring the information on the growth path of individual countries, being based on the pooling of observa-

---

and provides a biased picture.

[16]To disentangle these individual effects is not an easy task. Our procedure is the following. We start from the estimated Model 1:

$$STD_i = \hat{\beta}_0 + \hat{s}_{TGDP,TR}\left(TGDP_i, TR_i\right) + \hat{s}_{TGDP,TR,AS}\left(TGDP_i, TR_i, AS_i\right).$$

To identify the effect of e.g. AS we estimate the following equations

$$TGDP_i = \hat{s}_{AS}^{TGDP}\left(AS_i\right);$$
$$TR_i = \hat{s}_{AS}^{TR}\left(AS_i\right),$$

from which we obtained the fitted values $\overline{TGDP}_i$ and $\overline{TR}_i$. Finally we estimate the effect of AS on STD by:

$$\overline{STD}_i = \hat{\beta}_0 + \hat{s}_{TGDP,TR}\left(\overline{TGDP}_i, \overline{TR}_i\right) + \hat{s}_{TGDP,TR,AS}\left(\overline{TGDP}_i, \overline{TR}_i, AS_i\right).$$

The same procedure is repeated for $TGDP$ and TR.

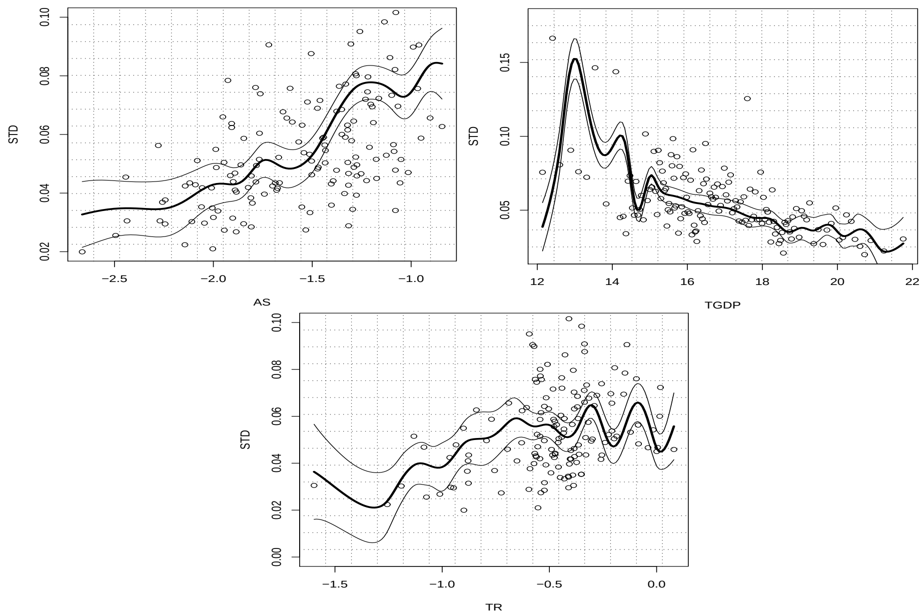

Figure 2: Estimation of *STD* for, respectively, log of *AS*, *TGDP*, and *TR*

tions and on the measurement of *GRV* by the standard deviation within a class. For example, consider two countries whose *GDP* belongs to the same class, having constant growth rates but at very different levels. If we compute the standard deviation of growth rates for that *GDP* class, we would obtain a high value, wrongly indicating high *GRV*. On the contrary, the method proposed in the next section, based on transition matrices, would correctly detect low volatility.[17]

4 Growth Volatility Indices

In this section we propose a set of synthetic indices to measure *GRV* and study their statistical properties. In particular, the measurement of *GRV* requires first the estimation of a Markov transition matrix, whose states $S = \{1, 2, ..., n\}$ represent growth rate classes. A transition matrix summarizes the information on the dynamics of growth rates (for more details see Quah (1993)), and is the basis to calculate *GRV* indices.

Heuristically, the indices quantify volatility by the *intensity of switches across growth rate classes*. The advantage of the approach based on transition matrices is that we can keep track of the dynamics of individual countries in the sample. To evaluate the relationship between *GRV* and, for

[17] Canning *et al.* (1998) avoid this specific problem by detrending data, but their procedure is not immune from introducing spurious volatility. At any rate we adopted their detrending procedure in Appendix C.

instance, *GDP* we calculate the values of these indices for different classes of *GDP*.

To define indices of *GRV* we draw on studies of inter- and intragenerational mobility of individuals (see, among others, Bartholomew (1982), pp. 24–30 and Shorrocks (1978)), and propose two new indices. Basically, these indices are functions of the elements of a transition matrix. In a transition matrix high values on the principal diagonal indicate low mobility, while the values of off-diagonal elements refer to changes of state and, therefore, high values of the latter are associated to high mobility.

A simple mobility index is the following, proposed by Shorrocks (1978):

$$I^S(\mathbf{P}) = \frac{n - trace(\mathbf{P})}{n-1}, \tag{12}$$

where $\mathbf{P}$ is a transition matrix of dimension n. The range of the index is $[0, n/(n-1)]$ and a high value means high mobility. However, I^S is not well-suited to measure growth volatility because it is not affected by the value of off-diagonal elements, a key point for the present analysis, but we refer to it as a term of comparison with the other indices discussed below.

Bartholomew (1982), p. 28, proposes the following index which takes explicitly into account the *distance* covered by a transition from i to j, $(i, j \in S)$, when the states correspond to increasing or decreasing values of a variable:

$$I^B(\mathbf{P}) = \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i p_{ij} \left| i - j \right|. \tag{13}$$

In I^B, p_{ij} is an element of the transition matrix $\mathbf{P}$, while π_i is an element of the associated ergodic distribution.[18] The range of I^B is $[0, 1]$: a higher value means higher mobility.

In this case only the absolute value of the difference between i and j is taken into account. It is worth verifying the effect of increasing the weight attached to "longer" jumps, in order to better appreciate the magnitude of the fluctuations. Therefore we introduce the following index:

$$I^{BM}(\mathbf{P}) = \frac{1}{(n-1)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i p_{ij} (i-j)^2, \tag{14}$$

in which the distance of the transition enters in a quadratic form. As before $I^{BM} \in [0, 1]$ and a higher value means higher mobility/volatility.

Indices I^B and I^{BM} weight the transitions from growth rate class i by the corresponding mass in the long-run equilibrium (i.e. in the ergodic distribution). In other words, considering the elements of the ergodic distribution

[18] The ergodic distribution represents the long-run distribution of the Markov process. For more details see Bartholomew (1982).

as weights amounts to measuring GRV in the long-run equilibrium. However, also the volatility along the transition path can reveal very interesting information. The following indices fill this gap:

$$I^{FL}(\mathbf{P}) = \frac{1}{A}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\,|i-j|\,; \tag{15}$$

$$I^{FLM}(\mathbf{P}) = \frac{1}{A^2}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}(i-j)^2. \tag{16}$$

I^{FL} and I^{FLM} respectively correspond to I^B and I^{BM}, except for the absence of the elements of the ergodic distribution. The constant A normalizes both indices to the range $[0,1]$.[19] A higher value still means higher mobility/volatility. In the next section we study the statistical properties of these indices.

4.1 Statistical Properties

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

The element p_{ij} of $\mathbf{P}$ represents the transition probability from state i to state j and therefore $\Sigma_{j=1}^{k} p_{ij} = 1$ and $0 \leq p_{ij} \leq 1$. Moreover, let p_i be the fraction of observations in initial state i, i.e. $p_i = \frac{n_i}{n}$.

Suppose the ergodic distribution for this process exists. then, the ergodic distribution is defined as

$$\pi = \pi \mathbf{P} \tag{17}$$

under the constraint

$$\pi \mathbf{u}' = 1,$$

where $\mathbf{u}$ is the sum vector. In the following we assume that the rows of $\mathbf{P}$ are independent.

4.1.1 Consistent Estimators

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

$$\hat{\mathbf{P}} = [\hat{p}_{ij}] = \left[\frac{n_{ij}}{n_i}\right], \tag{18}$$

[19] In particular:

$$A = \begin{cases} 2\Sigma_{i=\frac{n-1}{2}+1}^{n-1} i + \frac{n-1}{2} & \text{for } n \text{ odd;} \\ 2\Sigma_{i=\frac{n}{2}}^{n-1} i & \text{for } n \text{ even.} \end{cases}$$

where $n_i = \Sigma_{j=1}^{n} n_{ij}$ (for a proof see, e. g. Norris (1997), pp. 55-56). $\hat{\mathbf{P}}$ being the ML estimator, these estimates are consistent.

In general, take $\mathbf{P}$ and a function I such that $I : \mathbf{P} \rightarrow \Re$. Since $\mathbf{P}$ is unknown, then $I(\mathbf{P})$ is unknown as well. A natural estimator is $\hat{I} = I\left(\hat{\mathbf{P}}\right)$, which, in turn, is consistent (see Trede (1999)). I can represent any function (linear and non-linear), i.e. each index of *GRV* calculated on the basis of the transition matrix.

4.1.2 Distribution of Estimates

Stuart and Ord (1994), p. 260, show that the distribution of $\vec{n}_i$ converges to a n-variate normal distribution, with means $n_i p_{ij}$, variances $n_i p_{ij}(1 - p_{ij})$ and covariances $cov(n_{ij}, n_{iq}) = -n_i p_{ij} p_{iq}$. Thus $\sqrt{n_i}(\hat{p}_{ij} - p_{ij})$ tends towards the normal distribution $N(0; p_{ij}(1 - p_{ij}))$.

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

$$I\left(\hat{\mathbf{P}}\right) = I(\mathbf{P}) + DI(\mathbf{P})\left(\text{vec}\left(\hat{\mathbf{P}}' - \mathbf{P}'\right)\right),$$

where

$$DI(\mathbf{P}) = \frac{\partial I(\mathbf{P})}{\partial\, \text{vec}(\mathbf{P}')'} \tag{19}$$

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

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

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

where

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & & \\ & ... & \\ & & \mathbf{V}_k \end{bmatrix} \tag{20}$$

is a block diagonal with

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

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

Therefore the asymptotic distribution of I is given by:

$$\sqrt{n}\left(I\left(\hat{\mathbf{P}}\right) - I(\mathbf{P})\right) \xrightarrow{d} N\left(0, \sigma_I^2\right), \tag{21}$$

where

$$\sigma_I^2 = (DI(\mathbf{P}))\,\mathbf{V}\,(DI(\mathbf{P}))'. \tag{22}$$

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

$$\hat{\sigma}_I^2 = \left(DI\left(\hat{\mathbf{P}}\right)\right)\hat{\mathbf{V}}\left(DI\left(\hat{\mathbf{P}}\right)\right)'. \tag{23}$$

Since $I(\mathbf{P})$ is normally distributed, then $(1-\alpha)$-confidence interval for $I(\mathbf{P})$ is

$$I\left(\hat{\mathbf{P}}\right) \pm c\frac{\hat{\sigma}_I}{\sqrt{n}}, \tag{24}$$

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

$$s = \frac{I\left(\hat{\mathbf{P}}\right) - I(\mathbf{P})}{\frac{\hat{\sigma}_I}{\sqrt{n}}} \tag{25}$$

converges towards a Gaussian distribution under the null hypothesis $I\left(\hat{\mathbf{P}}\right) = I(\mathbf{P})$.

Finally, given two transition matrices $\hat{\mathbf{P}}^1$and $\hat{\mathbf{P}}^2$, we have that:

$$s = \frac{I\left(\hat{\mathbf{P}}^1\right) - I\left(\hat{\mathbf{P}}^2\right)}{\sqrt{\frac{\hat{\sigma}_{I^1}^2}{n} + \frac{\hat{\sigma}_{I^2}^2}{n}}}, \tag{26}$$

converges towards a Gaussian distribution under the null hypothesis $I\left(\hat{\mathbf{P}}^1\right) = I\left(\hat{\mathbf{P}}^2\right)$.

4.1.3 Analytical derivative of the ergodic distribution

DM provides a general procedure of testing. For our aim, a potential problem can arise for calculating the derivative of I with respect to elements of the ergodic distribution, in the computation of $DI(\mathbf{P})$ for volatility indices which include these elements. Conlisk (1985) provides an analytical formulation to tackle this problem. Assume that the increase in the element j in row i, p_{ij}, is absorbed by a decrease in the element of the last column k of row i, p_{ik} (the row sum must sum to one). Thus, the derivative of the $q-th$ element of the ergodic distribution is defined as follows:

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

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

5 Empirical Results

In the following we study the relation between *GRV*, *GDP*, *AS* and *TGDP* by calculating the values of the indices described in Section 4. As in Section 3, we first evaluate the individual effects of our variables, and then study their interactions, with particular attention to the explanatory power of *GDP*.

In particular, in a first stage: (i) we separate the observations on *GDP*, *AS* and *TGDP* in four classes for each variable, from "low" to "high" values; (ii) we calculate the transition matrix with five growth rate classes for each class, (iii) we compute indices (12), (13), (14), (15), (16) for every transition matrix and, finally, (iv) we make inference on these estimates. In a second stage we evaluate the interactions of the variables in this framework.

First we define the five growth rate classes common to all three variables. We set the central class to include the average growth rate of the sample, equal to 2%, and define the other classes symmetrically around this central class. With this criterion we obtain the state space:[20]

$$S = \{[-\infty, -2\%), [-2\%, 1\%), [1\%, 3\%), [3\%, 6\%), [6\%, +\infty)\}. \tag{27}$$

Alternatively, the state space for the calculation of the transition matrices could be based not on absolute values of growth rates, but on deviations from the trend.

5.1 Per Capita GDP

We define four *GDP* classes (in logs) which contain the same number of observations (≈ 1100), obtaining the following:

$$I = [0, 6.98), II = [6.98, 7.9), III = [7.9, 8.82), IV = [8.82, +\infty).$$

For every *GDP* class we estimate a transition matrix relative to the state space S.[21] Table 2 contains the values of the indices calculated for each of the four transition matrices. We observe that in all cases the value of the index is generally decreasing with respect to the *GDP* class and that, in particular, the value of the index in the first *GDP* class is always higher than in the last. This result broadly agrees with Figure 1 in which volatility is measured by the standard deviation of growth rates.

Table 3 reports the p-values of tests of a null hypothesis of equality between the value of the index in the first *GDP* class versus its value in each of the other *GDP* classes, for all the indices. Tests confirm that *GDP* class *I* generally has a statistically significant higher *GRV*. At a conventional

[20] Results are not affected by slight changes of the classes' limits.

[21] The four transition matrices and the four ergodic distributions (one for each *GDP* class) are obtainable with the codes available on the authors' websites.

Index\GDP class	I	II	III	IV
I^S	0.8313 (0.0195)	0.7831 (0.0192)	0.7417 (0.0193)	0.7793 (0.023)
I^B	0.2583 (0.0108)	0.2474 (0.0093)	0.2272 (0.0086)	0.1983 (0.0077)
I^{BM}	0.1548 (0.0099)	0.1382 (0.008)	0.1206 (0.0074)	0.0888 (0.0062)
I^{FL}	0.4178 (0.0149)	0.3658 (0.0126)	0.3136 (0.0115)	0.3253 (0.0162)
I^{FLM}	0.3199 (0.0169)	0.259 (0.0138)	0.2083 (0.0122)	0.2148 (0.0188)

Table 2: Growth volatility indices. Standard errors in parenthesis. *GDP*

Index\GDP class	I vs II	I vs III	I vs IV
I^S	0.04*	0*	0.04*
I^B	0.22	0.01*	0*
I^{BM}	0.10	0*	0*
I^{FL}	0*	0*	0*
I^{FLM}	0*	0*	0*

Table 3: Test of equality between the *GRV* index of *GDP* class I versus its value in the other classes. * means rejection of the null hypothesis of equality at 5% confidence level.

5% level, the null hypothesis is not rejected only in the comparison between the value of the index in the first and in the second *GDP* class for indices I^B and I^{BM} (but it is rejected at 10% for the latter).

To check if there is a monotonic decreasing relationship among the values of the indices at different *GDP* levels we also tested the following hypotheses of equality (details omitted): (i) *GDP* class II vs *GDP* class III; (ii) *GDP* class III vs IV. In case (i), the hypothesis is strongly rejected for I^{FL} and I^{FLM}, and is rejected at approximately 6% level for the other indices; in case (ii) the hypothesis is rejected only for I^B and I^{BM}. However, in the other cases we do not reject the null hypothesis that the indices in *GDP* classes III vs IV are equal (note that in Table 2 the value of the index in *GDP* class IV is actually higher than in *GDP* class III for I^S, I^{FL} and I^{FLM}). Hence, according to indices I^B and I^{BM}, the decrease is statistically significant when we move from class II onwards, while with I^{FL} and I^{FLM} the decrease is statistically significant from class I, but for the two higher *GDP* classes the relation may be flat. Also, for I^S we do not find evidence of a monotonic decrease as the value of the index in *GDP* classes III and IV may be equal.

Overall, the indices indicate the presence of a negative relationship between *GRV* and *GDP*, which may become flat in some *GDP* ranges. In any case, a comparison between the first and the last *GDP* classes always shows a significantly higher volatility in the former.

5.2 Structural Change

In this section we address the relationship between *GRV* and structural change proxied by *AS*. We first define four *AS* classes with the same number of observations (≈ 784). The resulting classes' limits are:

$$I = [0, 0.08), II = [0.08, 0.2), III = [0.2, 0.33), IV = [0.33, 1].$$

Table 4 contains the volatility indices calculated with this class definition. Results seems to be in accordance with the pattern in Figure 2. Moving from high to low levels of *AS*, that is following a typical development path, volatility decreases from class IV to class III, then increases in class II and decreases again in class I. However, tests of equality between the value of the indices in class III and in class II do not allow to reject the null hypothesis at conventional 5% level.[22]. Finally, volatility is significantly higher in class IV than in class I: the hypothesis of equality between the value of the index in class I and in class IV is strongly rejected for all indices (we omit the details of the tests).

At this stage, we take this result as indicating the possible presence of a more complex behavior at intermediate levels of *AS*, which is in accordance with the non–monotonic pattern of *STD* found in Figure 1. Notice that indices calculated for classes II and III are not significantly different at 5% level, but only at about 15%.

5.3 The Dimension of the Economy

In this section we repeat the exercise considering *TGDP* to proxy for the dimension of the economy.[23] We define four *TGDP* classes (in logs) with

[22]The p-values of the tests for I^S, I^B, I^{BM}, I^{FL}, I^{FLM} are, respectively, 0.15, 0.14, 0.17, 0.10, 0.13.

[23]In the next section we examine the interaction of *TGDP* with *TR*.

Index\AS class	I	II	III	IV
I^S	0.7428 (0.0229)	0.8152 (0.0214)	0.7834 (0.0223)	0.8756 (0.0207)
I^B	0.2297 (0.0088)	0.2735 (0.0105)	0.2576 (0.0102)	0.3199 (0.0115)
I^{BM}	0.1105 (0.007)	0.1501 (0.0094)	0.1377 (0.0092)	0.1915 (0.0114)
I^{FL}	0.3037 (0.0132)	0.3611 (0.0133)	0.3364 (0.0136)	0.4119 (0.0138)
I^{FLM}	0.1905 (0.0134)	0.2478 (0.014)	0.2249 (0.0142)	0.2975 (0.0151)

Table 4: Growth volatility indices. Standard errors in parenthesis. *AS*

the same number of observations (≈ 1100):

$$I = [0, 0.03), II = [0.03, 0.05), III = [0.05, 0.08), IV = [0.33, 1].$$

With this class definition, and with the same state space for growth rate classes, we obtain the volatility indices in Table 5. In this case we observe a monotonic decrease for indices I^B, I^{BM} and I^{FL} across the *TGDP* classes, while for I^{FLM} the value is higher in *TGDP* class III than in II. Finally, no clear relation emerges from I^S.

From Table 6 we see that, with the exception of I^S, the value of the index in *TGDP* class I is, in most cases, significantly higher than the values in other classes. However, (i) in neither case we can reject the hypothesis of equality between the values in *TGDP* classes II and III (details omitted), (ii) we always reject the hypothesis of equality between the values of the indices in classes III and IV. This is in agreement with Figure 1, in which the relation between *STD* and *TGDP* is slightly flatter at intermediate *TGDP* levels.

Again, we find a broad confirmation of the existence of a negative relation between *STD* and *TGDP*. As for the case of *AS*, we find a clearer negative relation when we compare the values of the indices in the extreme classes, while the relation appears flatter at intermediate levels.

5.4 On Conditioning

In Section 3 we reported the results of nonparametric estimations suggesting that *GDP* is not informative when *TGDP*, *TR* and *AS* are considered. In other words, when the latter explanatory variables are present in a regression, *GDP* does not provide further information on *GRV*.

Here we address this issue in the approach based on the Markov transition matrix. First, notice that the analysis in the previous section can be considered as deriving from the estimation of *conditioned* transition matrices. In fact, the basis for the calculation of each single index is a transition matrix indicating the probabilities to observe transitions across growth rate

Index\TGDP class	I	II	III	IV
I^S	0.7721 (0.0185)	0.7998 (0.0182)	0.7799 (0.0186)	0.7339 (0.0193)
I^B	0.2904 (0.0104)	0.2749 (0.0085)	0.2645 (0.0088)	0.2155 (0.0072)
I^{BM}	0.18 (0.01)	0.1524 (0.0076)	0.1462 (0.0081)	0.1001 (0.0055)
I^{FL}	0.3682 (0.0116)	0.3478 (0.0107)	0.3449 (0.0115)	0.2846 (0.0103)
I^{FLM}	0.2725 (0.0124)	0.2326 (0.011)	0.2375 (0.0121)	0.1679 (0.01)

Table 5: Growth volatility indices. Standard errors in parenthesis. *TGDP*

Index\TGDP class	I vs II	I vs III	I vs IV
I^S	0.14	0.38	0.08
I^B	0.12	0.03*	0*
I^{BM}	0.01*	0*	0*
I^{FL}	0.10	0.07	0*
I^{FLM}	0*	0.02*	0*

Table 6: Test of equality between the *GRV* index of *TGDP* class I versus its value in the other classes. * means rejection of the null hypothesis of equality at 5% confidence level.

classes, starting from a given growth rate class *and* a given e.g. *TGDP* class.[24]

The formal definition of a transition probability from growth rate class S_i to growth rate class S_j, given that the observation is in *TGDP* class I is:

$$p\left(g_t \in S_j | g_{t-1} \in S_i, TGDP_{t-1} \in I\right) = \\ = p\left(g_t \in S_j | g_{t-1} \in S_i\right) \left[\frac{p\left(TGDP_{t-1} \in I | g_t \in S_j, g_{t-1} \in S_i\right)}{p\left(TGDP_{t-1} \in I | g_{t-1} \in S_i\right)}\right]. \tag{28}$$

In Equation (28) the term on the left–hand side is an element of the conditioned transition matrix from which we derived our indices relative to *TGDP* class I. The first term on the right-hand side is an element of the unconditioned transition matrix for growth rates, and the second term reflects the probability that that the transition starts from a state where the growth rate is associated to a *TGDP* in class I.

If the conditioning variable *TGDP* is not relevant, $p\left(g_t \in S_j | g_{t-1} \in S_i, TGDP_{t-1} \in I\right) = p\left(g_t \in S_j | g_{t-1} \in S_i\right)$: any transition matrix calculated considering alternative values of *TGDP* would not be statistically different from the unconditioned transition matrix.[25] Therefore, *GRV* indices calculated from the former would not be statistically different from each other, and from those calculated from the unconditioned transition matrix. In the same manner, if we condition on two variables, e.g. *TGDP* and *GDP*, we have:

$$p\left(g_t \in S_j | g_{t-1} \in S_i, TGDP_{t-1} \in I, GDP_{t-1} \in I\right) = \\ = p\left(g_t \in S_j | g_{t-1} \in S_i, TGDP_{t-1} \in I\right) * \\ * \left[\frac{p\left(GDP_{t-1} \in I | g_t \in S_j, g_{t-1} \in S_i, TGDP_{t-1} \in I\right)}{p\left(GDP_{t-1} \in I | g_{t-1} \in S_i, TGDP_{t-1} \in I\right)}\right] \tag{29}$$

[24] We could have estimated an *unconditioned* transition matrix for growth rate classes only, which does not distinguish among the *TGDP* levels associated to each transition. An example of conditioned Markov chains is in Quah (1996).

[25] From the condition $p\left(g_t \in S_j | g_{t-1} \in S_i, TGDP_{t-1} \in I\right) = p\left(g_t \in S_j | g_{t-1} \in S_i\right)$ derives $p\left(TGDP_{t-1} \in I | g_t \in S_j, g_{t-1} \in S_i\right) = p\left(TGDP_{t-1} \in I | g_{t-1} \in S_i\right)$, i.e. the information on $TGDP_{t-1}$ is irrelevant to know the state of g_t.

and the same reasoning for the relevance of *GDP*, given *TGDP*, applies.

Here we do not provide a complete discussion of this issue, but only some evidence on the relevance of *TR*, *AS* and *GDP* in explaining *GRV* given *TGDP*. Namely, we compare the values of two indices, I^B and I^{FL}, computed for *TGDP* classes only, with the values obtainable when the transition matrix is calculated for every *TGDP* class conditioned to each class of *AS*, *TR* and *GDP*. Clearly, results are not completely comparable with those of Section 3. In that case more variables were considered jointly, while here we investigate pairwise relations, in which there is one principle variable, *TGDP* or *AS*, and only another one interacting with it.

Tables 7 and 8 consider the relevance of the information provided by *AS* when *TGDP* is the principal variable, respectively, for index I^B and index I^{FL}.

The first column of the tables contains the volatility index obtained for *TGDP* classes (see Table 5). The other columns contain the values of the index when, for each *TGDP* class, we condition on each *AS* class. If *AS* matters, then *GRV* should increase with *AS*, and the conditioned indices should be statistically different from the unconditioned indices in the first column.[26]

What we observe is that: given a *TGDP* level, an increase in *AS* is generally associated to an increase in *GRV* (at least if we compare *AS(I)* and *AS(IV)*), confirming the insight that a higher agricultural share causes a higher *GRV*. However, the conditioned indices seem to be statistically different from the unconditioned ones especially for *TGDP(I)*. In fact, for both I^B and I^{FL}, in three out of four cases the difference is statistically significant at 5% level.

Tables 9 and 10 analyse the relation between *TGDP* and *TR*.

[26]Each conditioned indices is calculated starting from the observations belonging to a *TGDP* class and an *AS* class. We chose to consider only indices calculated from a minimum number of observations, at least 75. An alternative test can be conducted, aiming at testing the *joint* significance of the differences between the conditioned and unconditioned indices. We leave this issue on a side for future research.

	Uncond.	$AS(I)$	$AS(II)$	$AS(III)$	$AS(IV)$
$TGDP(I)$	0.2904 (0.0104)	0.1923** (0.02921)	0.3079 (0.02720)	0.3510 ** (0.02908)	0.3419** (0.01852)
$TGDP(II)$	0.2749 (0.0085)	0.3174** (0.02348)	0.2725 (0.02304)	0.2503 (0.01817)	0.2958 (0.01962)
$TGDP(III)$	0.2645 (0.0088)	0.2517 (0.02355)	0.2989** (0.01961)	0.2188* (0.01912)	0.2876 (0.03172)
$TGDP(IV)$	0.2155 (0.0072)	0.2027 (0.01091)	0.2204 (0.01741)	0.2307 (0.01859)	–

Table 7: Values of I^B for *TGDP* conditioned on *AS*. ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

	Uncond.	$AS(I)$	$AS(II)$	$AS(III)$	$AS(IV)$
$TGDP(I)$	0.3682 (0.0116)	0.2650** (0.03750)	0.3812 (0.03108)	0.4679** (0.03331)	0.4377** (0.02014)
$TGDP(II)$	0.3478 (0.0107)	0.3935* (0.02679)	0.3641 (0.02966)	0.3532 (0.02738)	0.4010** (0.02570)
$TGDP(III)$	0.3449 (0.0115)	0.3216 (0.02692)	0.4099 (0.02415)	0.3037* (0.02582)	0.3932 (0.03876)
$TGDP(IV)$	0.2846 (0.0103)	0.3330* (0.02834)	0.2929 (0.02299)	0.2984 (0.02618)	-

Table 8: Values of I^{FL} for $TGDP$ conditioned on AS. ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

	Uncond.	$TR(I)$	$TR(II)$	$TR(III)$	$TR(IV)$
$TGDP(I)$	0.2904 (0.0104)	0.3494* (0.03658)	0.3643** (0.03172)	0.2838 (0.02427)	0.3080 (0.01845)
$TGDP(II)$	0.2749 (0.0085)	0.2875 (0.02191)	0.2668 (0.02064)	0.2810 (0.01773)	0.2900 (0.01771)
$TGDP(III)$	0.2645 (0.0088)	0.2901 (0.02303)	0.2581 (0.01893)	0.2363* (0.01873)	0.2399 (0.01828)
$TGDP(IV)$	0.2155 (0.0072)	0.2404** (0.01331)	0.2000 (0.01275)	0.1936* (0.01453)	0.1825* (0.01903)

Table 9: Values of I^B for $TGDP$ conditioned on TR. ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

	Uncond.	$TR(I)$	$TR(II)$	$TR(III)$	$TR(IV)$
$TGDP(I)$	0.3682 (0.0116)	0.4671** (0.03962)	0.4684** (0.03220)	0.3691 (0.02738)	0.3871 (0.02101)
$TGDP(II)$	0.3478 (0.0107)	0.3817 (0.02819)	0.3600 (0.02774)	0.4064** (0.02600)	0.3632 (0.02177)
$TGDP(III)$	0.3449 (0.0115)	0.3678 (0.02748)	0.3551 (0.02418)	0.3375 (0.02755)	0.3120* (0.02231)
$TGDP(IV)$	0.2846 (0.0103)	0.3380** (0.01889)	0.2722 (0.02386)	0.2578 (0.02199)	0.2359** (0.02489)

Table 10: Values of I^{FL} for $TGDP$ conditioned on TR. ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

We expect that, given *TGDP,* an increase in *TR* is associated to a decrease in *GRV*. Indeed, we find that, comparing the values of the indices for *TR(I)* and *TR(IV)*, *GRV* generally decreases. However, the conditioned indices are statistically different from the unconditioned ones especially for *TGDP(I)* and *TGDP(IV).* At 5% or 10% level, the difference is statistically significant for two or three *TR* classes.

5.4.1 Conditioning on GDP

We concentrate now on the role of *GDP*. Our hypothesis is that the information on *GDP* is not relevant when we control for the dimension of the economy and for structural change, proxied by the share of the agricultural sector. We first evaluate *GDP* against *TGDP* in Tables 11 and 12.

We expect to find a decreasing *GRV* with *GDP*, and this tendency can be partially found in the results.[27] On the other hand, there appear to be no *TGDP* class for which the inclusion of GDP produces statistically different values for the *GRV* indices, with the exception of *TGDP(III)* for I^B.

However, the most important test for the relevance of *GDP* is the analysis of its role in presence of *AS*, which is directly connected to economic development. Tables 13 and 14 contain the results when the principal variable is *AS* and we condition on *GDP*.

Results are in Tables 13 and 14 (values of the unconditioned indices in the first column are from Table 4). First of all notice that many values are not available for lack of data. This was predictable as it is likely to have very few observations for, say, *AS(I)* and *GDP(I).* This can be a first hint on the irrelevance of conditioning on *GDP* in presence of *AS*. Moreover, the number of statistically significant differences between the unconditioned and conditioned values is particularly low, if compared with the previous cases.

Summing up: we have attempted to identify the relative role of our vari-

[27] The value of 0.4427 in Table 12 is based on only 88 observations and is therefore scarcely relevant.

	Uncond.	$GDP(I)$	$GDP(II)$	$GDP(III)$	$GDP(IV)$
$TGDP(I)$	0.2904 (0.0104)	0.3197* (0.01579)	0.2956 (0.01855)	0.2493** (0.02079)	–
$TGDP(II)$	0.2749 (0.0085)	0.2763 (0.01673)	0.2706 (0.01468)	0.2559 (0.01424)	0.3272 (0.031413)
$TGDP(III)$	0.2645 (0.0088)	0.3214** (0.02679)	0.2661 (0.01948)	0.2414* (0.01520)	0.2378* (0.014905)
$TGDP(IV)$	0.2155 (0.0072)	–	0.2183 (0.02001)	0.2275 (0.01546)	0.1949** (0.008536)

Table 11: Values of I^B for *TGDP* conditioned on *GDP.* ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

	Uncond.	$GDP(I)$	$GDP(II)$	$GDP(III)$	$GDP(IV)$
$TGDP(I)$	0.3682 (0.0116)	0.4102** (0.01731)	0.3770 (0.02158)	0.3435 (0.02594)	–
$TGDP(II)$	0.3478 (0.0107)	0.3939** (0.02278)	0.3584 (0.01889)	0.3257 (0.01936)	0.4427** (0.04050)
$TGDP(III)$	0.3449 (0.0115)	0.4263 (0.03413)	0.3507 (0.02367)	0.3074* (0.01851)	0.3173 (0.02052)
$TGDP(IV)$	0.2846 (0.0103)	–	0.3067 (0.03041)	0.2891 (0.02013)	0.2900 (0.01776)

Table 12: Values of I^{FL} for *TGDP* conditioned on *GDP*. ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

	Non cond.	$GDP(I)$	$GDP(II)$	$GDP(III)$	$GDP(IV)$
$AS(I)$	0.2297 (0.0088)	–	–	0.2464 (0.01646)	0.2258 (0.01071)
$AS(II)$	0.2735 (0.0105)	–	0.3010 (0.02414)	0.2583 (0.01363)	0.2346** (0.02073)
$AS(III)$	0.2576 (0.0102)	0.2965* (0.02742)	0.2480 (0.013)	0.2504 (0.02061)	–
$AS(IV)$	0.3199 (0.0115)	0.3260 (0.01340)	0.3091 (0.02383)	–	–

Table 13: Values of I^{B} for *AS* conditioned on *GDP*. ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

	Non cond.	$GDP(I)$	$GDP(II)$	$GDP(III)$	$GDP(IV)$
$AS(I)$	0.3037 (0.0132)	–	–	0.3133 (0.02310)	0.3103 (0.01724)
$AS(II)$	0.3611 (0.0133)	–	0.4176** (0.02728)	0.3349 (0.01752)	0.3327 (0.02742)
$AS(III)$	0.3364 (0.0136)	0.3954* (0.03469)	0.3213 (0.02367)	0.3350 (0.01851)	–
$AS(IV)$	0.4119 (0.0138)	0.4246 (0.01615)	0.3858 (0.02735)	–	–

Table 14: Values of I^{FL} for *AS* conditioned on *GDP*. ** indicates rejection of null hypothesis of equality to the index in the first column at 5% level; * at 10% level

ables in the explanation of *GRV* in the approach based on transition matrices. At this stage, we have found a partial confirmation of the hypotheses formulated from the model in Section 2, on the relevance of the dimension of the economy and structural change and on the irrelevance of per capita GDP in the explanation of growth volatility. In addition, in the analysis of conditioned *GRV* indices, clearer results appear more often at extreme *TGDP* classes, indicating that in the transition from low to high *TGDP* levels, the relations are more blurred (as resulted also from the preliminary graphical analysis).

6 Conclusions

This paper investigates the relation between growth volatility and the level of development, structural change and the size of the economy. Two methods used to measure growth volatility, (i) the standard deviation of the growth rate and (ii) a set of indices inspired by the literature on social mobility, substantially lead to the same results. Growth volatility appears to be negatively related to total GDP, proxy for the dimension of the economy. In particular it seems appropriate to consider as an additional control for the dimension of the economy the integration in the world markets. Moreover, growth volatility appears to be negatively related to the share of agriculture on GDP, proxy for structural change. Finally, per capita GDP, proxing for the level of development, does not seem to add relevant information when the other variables are considered. A direction for further research may be an assessment of the explanatory power of other factors related to development and to growth volatility, like the growth of a financial sector, in relation to structural change.

References

Acemoglu, D. and F. Zilibotti (1997), "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth", *Journal of Political Economy*, 105, pp. 709-751.

Acemoglu, D., S. Johnson, J. Robinson and Y. Thaicharoen (2003), "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth", *Journal of Monetary Economics*, 50, pp. 49-123.

Bartholomew, D.J. (1982), *Stochastic Models for Social Processes*, 3rd edition, John Wiley & Son, New York.

Bowman, A.W and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis*, Oxford, Clarendon Press

Canning, D., L. Amaral, Y. Lee, M. Meyer and H. Stanley (1998) "Scaling the Volatility of GDP Growth Rate", *Economics Letters*, 60, 335-341.

Conlisk J. (1985). "Comparative Statics for Markov Chains". *Journal of Economic Dynamics and Control*, 9, pp. 139-151.

Grossman, H. and M. Kim (1996), "Predation and Accumulation", *Journal of Economic Growth*, 1(3), 333-350.

Horvarth, M. (1998), "Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks", *Review of Economic Dynamics*, 1, 781-808.

Maddison A. (2001), *The World Economy: a Millenium Prospective.* Paris: OECD.

Norris, J. (1997), *Markov Chains*, Cambridge University Press, Cambridge.

Pasinetti, L. (1981), *Structural Change and Economic Growth : a Theoretical Essay on the Dynamics of the Wealth of Nations*, Cambridge University Press, Cambridge.

Pritchett, L. (2000), "Understanding Patters of Economic Growth: Searching for Hills among Plateaus, Mountains, and Plains", *The World Bank Economic Review*, 14, 221-250.

Quah D. T. (1993), "Empirical Cross-section Dynamics for Economic Growth", *European Economic Review*, 37, 426-434.

Quah D. T. (1996), "Regional Convergence Clusters Across Europe", *European Economic Review*, 40, 951-958.

Ramey, G. and Ramey, V. A. (1995), "Cross-country Evidence on the Link Between Volatility and Growth", *American Economic Review*, 85, 1138-1151.

Scheinkman J. and M. Woodford (1994), "Self organized criticality and economic fluctuations". *American Economic Review*, 84, 417-421.

Shorrocks, A. F. (1978), "The Measurement of Mobility", *Econometrica*, 46, 1013-1024.

Easterly, W., R. Islam and J. E. Stiglitz (2000), "Explaining Growth Volatility", mimeo

Stuart A. and K. Ord (1994). *Kendall's Advanced Theory of Statistics*, 6th ed., vol. I, Edward Arnold, London.

Trede M. (1999). "Statistical Inference for Measures of Income Mobility", *Jahrbücher fur Nationalökonomie und Statistik*, March, v. 218(3-4), pp. 473-490.

Wood S. N. (2000). "Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties", *Journal of the Royal Statistical Society B*, 62, pp. 413-428.

A Country List

B GAM estimation with cross-country data

In this Appendix we show the results of GAM estimations with cross-country data, restricting our analysis to the period $1970-1998$ for lack of data on TR and AS. For each country we consider the standard deviation of growth rates for the period as STD, the value of per capita GDP in 1970 as GDP, the value of total GDP in 1970 for $TGDP$ and the average value of trade openness and the share of agriculture on GDP for the period $1970-1975$ as TR and AS (possible missing values have been removed). The available observations are only 87 (we would have only 58 observations if 1960 were the initial year). Table 16 reports the results of GAM estimations.

Results for Model 1 are not reported because the routine for the likelihood minimization could not reach convergence. From the other models it results that TR is not significant, but GDP is and Model 6 is the best in terms of GCV score. However, we remark that lack of observations is particularly notable for low-income countries, and this could bias these results.

C GAM estimation based on panel regression

Here we measure GRV by means of the standard deviation of residuals from a panel regression of growth rates against a common component for each period and a country fixed effect, as in Canning *et al.* (1998).[28] As above we calculate the residuals for all observations; after pooling and partitioning the latter into classes on the basis of $TGDP$, we calculate the standard deviation of the residuals for each class. Table 17 reports the results of estimations.

We see again that the inclusion of GDP in Model 6 is not significant. However, we do not find a clear evidence that Model 1 is the best specification. Model 2 provides comparable results, that is TR could be not relevant for explaining GRV.

D GAM estimation with deviation from an autoregressive process

In this appendix we measure the volatility by the innovation from a second-order autoregressive process for growth rates, as suggested by Acemoglu and Zilibotti (1997), but we use first differences of growth rates, in the light of

[28] In particular we estimate the following panel:

$$g_{it} = \mu + \phi_t + \delta_i + \varepsilon_{it},$$

where g_{kt} is the growth rate at time t of country k and ε_{kt} are the residuals to be calculated.

AFRICA	1 Algeria	2 Angola	3 Benin	4 Botswana
5 Cameroon	6 Cape Verde	7 Cent. Afr. Rep.	8 Chad	9 Comoros
10 Congo	11 Côte d' Ivoire	12 Djibouti	13 Egypt	14 Gabon
15 Gambia	16 Ghana	17 Kenya	18 Liberia	19 Madagascar
20 Mali	21 Mauritania	22 Mauritius	23 Morocco	24 Mozambique
25 Namibia	26 Niger	27 Nigeria	28 Rwanda	29 Senegal
30 Seychelles	31 Sierra Leone	32 Somalia	33 South Africa	34 Sudan
35 Swaziland	36 Tanzania	37 Togo	38 Tunisia	39 Uganda
40 Zambia	41 Zimbabwe	LATIN AMERICA	42 Argentina	43 Brazil
44 Chile	45 Colombia	46 Mexico	47 Peru	48 Uruguay
49 Venezuela	50 Bolivia	51 Costa Rica	52 Cuba	53 Dominican Rep.
54 Ecuador	55 El Salvador	56 Guatemala	57 Haiti	58 Honduras
59 Jamaica	60 Nicaragua	61 Panama	62 Paraguay	63 Puerto Rico
64 Trin. Tobago	OFF WESTERN	65 Australia	66 New Zealand	67 Canada
68 United States	WEST ASIA	69 Bahrain	70 Iran	71 Iraq
72 Israel	73 Jordan	74 Kuwait	75 Lebanon	76 Oman
77 Qatar	78 Saudi Arabia	79 Syria	80 Turkey	81 UAE
82 Yemen	83 W.Bank Gaza	EAST ASIA	84 China	85 India
86 Indonesia	87 Japan	88 Philippines	89 South Korea	90 Thailand
91 Bangladesh	92 Hong Kong	93 Malaysia	94 Nepal	95 Pakistan
96 Singapore	97 Sri Lanka	98 Afghanistan	99 Cambodia	100 Laos
101 Mongolia	102 North Korea	103 Vietnam	EUROPE	104 Austria
105 Belgium	106 Denmark	107 Finland	108 France	109 Germany
110 Italy	111 Netherlands	112 Norway	113 Sweden	114 Switzerland
115 UK	116 Ireland	117 Greece	118 Portugal	119 Spain

Table 15: Country list

Model	**1**	**2**	**3**	**4**	**5**	**6**
Constant	-	0	0	0	0	0
$s(TR, TGDP)$	-	-	0	-	-	-
$s(AS, TGDP)$	-	0.08	-	-	-	-
$s(AS, TR, TGDP)$	-	-	-	-	0	-
$s(TGDP)$	-	0	-	0	-	0
$s(AS)$	-	-	-	0	-	0.047
$s(TR)$	-	-	-	0.147	-	-
$s(GDP)$	-	-	-	-	-	0
GCV score($*10^{-4}$)	-	5.2677	5.5128	4.5165	5.138	3.876
R^2	-	0.302	0.217	0.484	0.842	0.601
Number of obs.	-	87	87	87	87	87

Table 16: Estimation of Equation (10). Dependent variable is *STD*. The p-value of the explanatory variables is reported

Model	**1**	**2**	**3**	**4**	**5**	**6**
Constant	0	0	0	0	0	0
$s(TR, TGDP)$	0	-	0	-	-	0
$s(AS, TGDP)$	-	0	-	-	-	-
$s(AS, TR, TGDP)$	0.013	-	-	-	0	0.01
$s(TGDP)$	-	0	-	0	-	-
$s(AS)$	-	-	-	0.15	-	-
$s(TR)$	-	-	-	0.59	-	-
$s(GDP)$	-	-	-	-	-	0.29
GCV score(*10^{-4})	3.378	3.3472	3.4398	3.451	3.641	3.394
R^2	0.627	0.578	0.349	0.401	0.375	0.646
Number of obs.	149	149	149	149	149	149

Table 17: Estimation of Equation (10). Dependent variable is STD, estimated by the residuals of a panel regression. The p-value of the explanatory variables and the estimated degrees of freedom (in parenthesis) are reported

the non-stationarity of most of countries' growth process (only 38 out of 119 countries pass the ADF test at 10% level). As above we calculate the residuals for all observations; after pooling and partitioning the latter into classes on the base of $TGDP$, we calculate the standard error of the residuals for each class. Table 18 reports the results of estimations.

Here Model 2 is the best specification, that is TR is not relevant to explain GRV. We see again that the inclusion of GDP in Model 6 (based on the Model 2) is not significant.

Model	**1**	**2**	**3**	**4**	**5**	**6**
Constant	0	0	0	0	0	0
$s(TR, TGDP)$	0.025	-	0	-	-	-
$s(AS, TGDP)$	-	0	-	-	-	0
$s(AS, TR, TGDP)$	0.583	-	-	-	0	-
$s(TGDP)$	-	0	-	0	-	0
$s(AS)$	-	-	-	0.58	-	-
$s(TR)$	-	-	-	0.27	-	-
$s(GDP)$	-	-	-	-	-	0.93
GCV score($*10^{-4}$)	4.6371	4.1325	4.4941	4.5448	4.652	4.2751
R^2	0.511	0.613	0.371	0.372	0.518	0.554
Number of obs.	149	149	149	149	149	149

Table 18: Estimation of Equation (10). Dependent variable is STD, estimated by the residuals of a autoregressive model. The p-value of the explanatory variables and the estimated degrees of freedom (in parenthesis) are reported