Panel Growth Regressions with General Predetermined Variables: Likelihood-Based Estimation and Bayesian Averaging

Enrique Moral-Benito*

CEMFI

January 2009

PRELIMINARY AND INCOMPLETE

ABSTRACT

Empirical growth regressions are plagued by biases and inconsistencies from an econometric perspective. To address these issues, likelihood-based estimation in a dynamic panel data with feedback model is introduced. This estimator can be interpreted as the LIML counterpart for standard GMM estimators of panel data models with general endogenous or predetermined variables and fixed effects. Finally, the estimator together with model averaging techniques are applied to empirical panel growth regressions. The results indicate that the estimated convergence rate is not significantly different from zero. Moreover, there seems to be only one variable, the investment ratio, that causes economic growth.

*E-mail address: enrique.moral@gmail.com. I wish to thank Manuel Arellano for constant guidance and advice.

1 Introduction

Over the last three centuries the growth of nations has generated worries for economists all around the world. Therefore, it has been studied continually since the days of Adam Smith. Nowadays, there exists a better understanding of the sources of economic growth. But the subject has proved elusive, and many unsolved problems and open questions remain.

As pointed out by Durlauf et al. (2005), the stylized facts of economic growth have led to two major issues in the development of formal econometric analysis of growth. The first one revolves around the question of convergence: are contemporary differences in growth rates across countries transient over sufficiently long time horizons? The second issue concerns the identification of growth determinants: which factors seem to explain observed differences in aggregate economies?

These two questions have been addressed by a huge literature on empirical growth regressions. However, this industry is plagued by econometric inconsistencies. These econometric problems arise not only when estimating an empirical growth model but also when selecting that model.

The second problem is known as model uncertainty and it emerges because theory does not provide enough guidance to select the proper empirical model. Model averaging techniques construct parameter estimates that formally address the dependence of model-specific estimates on a given model. There seems to be consensus about model averaging as the most promising solution to model uncertainty.

However, the first problem is still unsolved. Problems with estimating an empirical growth model are well known. The right-hand side variables are typically endogenous and measured with error. Omitted variable bias also arises because of the presence of unobservable time-invariant country-specific characteristics correlated with one or more of the regressors. The most prominent way to address these problems is the use of panel data econometric techniques that allow including country-specific fixed effects in the empirical model[1]. In particular, first-differenced generalized method of moments estimators applied to dynamic panel data models has been the most promising econometric method in empirical growth research. This kind of estimation procedure addresses the question of correlated individual effects and the issue of endogeneity and it was first proposed by Holtz-Eakin et al. (1988) and Arellano and Bond (1991). In the growth literature, this approach was first introduced by Caselli et al. (1996). Despite of its important advantages over simple cross-section regressions and other estimation methods for dynamic panel data models, it is now well known that in the growth context this method suffers from large finite sample biases. Given the variables considered in empirical growth models, the time series are

[1]Typical growth panels are based on a sample of N countries observed over ten or five-year periods. Despite some exercesises are carried out with five-year periods, I focus here on ten-year periods since we are interested in long-run economic growth.

persistent and the number of observations is typically small. Under these conditions, the first-differenced GMM estimator is poorly behaved. The reason is that lagged levels of the variables are only weak instruments for subsequent first-differences. This weak instruments problem may be present in other situations with highly persistent data in a small-T panel setting. In particular, microeconometric panels such as PSID, usually include persistent variables such as wages. If we are willing to avoid stationarity assumptions[2], as we are in the growth context, there is no better alternative proposed for this situation.

Against this background, this paper presents a feasible likelihood-based estimator in a panel data context equivalent to one-step first-differenced GMM in terms of the assumptions required, but that it does not suffer from weak instruments. Therefore, it provides consistent and unbiased estimates in the context of empirical growth regressions. After concentrating the resultant log-likelihood, the estimator is easy to apply by means of numerical optimization. Moreover, in the context of empirical growth regressions I also take into account model uncertainty by employing the Bayesian Averaging of Maximum Likelihood Estimates (BAMLE) proposed by Moral-Benito (2007).

I also argue that the estimator can be applied to a broad range of situations. For instance in the estimation of production functions we typically face two problems: (i) the regressors (employment and stock of capital) are potentially correlated with firm-specific fixed effects and productivity shocks, and, (ii) both employment and capital are highly persistent processes. Therefore, first-differenced GMM has poor finite sample properties in this context. Some authors have proposed to incorporate stationarity assumptions to the model and employ the denominated system-GMM estimator in order to alleviate the weak instruments problem (see for example Blundell and Bond (2000)). However, as in the growth context, the likelihood-based estimator proposed in this paper is able to solve the weak instruments problem present in the estimation of production functions without making any additional assumption. By the same token, there are many other situations in which the econometric issues just described are also present. As long as we are willing to avoid auxiliary assumptions about the stationarity of the series, I argue that the finite sample properties of the likelihood-based estimator introduced in the present paper are better than those of other available estimators.

Taking up again the growth framework, closely related to the econometric issues mentioned above, there is the convergence debate. After two decades of research the question is still unanswered: Is there conditional convergence across countries? Some authors consider that the available empirical evidence supports the conditional convergence hypothesis predicted by the neoclassical growth model. However, from a skeptical point of view, the lack of reliable estimates of the convergence parameter in growth regressions is enough to hamper consensus on the answer of this relevant question.

On the other hand, despite some progress has been done, there is no clear evidence on

[2] Assuming mean stationarity of the variables, we can exploit additional moment conditions and employ the so-called system-GMM estimator as proposed in Arellano and Bover (1995).

the most prominent variables in fostering economic growth. The reason is that all previous studies attempting to solve this issue, may be subject to criticism because they suffer from econometric inconsistencies and biases.

Given the above, after considering all potential sources of biases and inconsistencies, I obtain that conditional convergence is not present across the countries in my sample. In particular the estimated speed of convergence is 0.73%, but looking at the standard error, it is not significantly different from zero. This result would lead us to conclude that the hypothesis of no conditional convergence can not be rejected given the available data.

On the other hand, in contrast to previous consensus in the literature, I also conclude that there is only one variable that seems to robustly cause economic growth, the investment ratio. This conclusion is derived from the fact that my approach is the first one in the empirical growth literature that allows obtaining unbiased estimates of what can be interpreted as causal effects. Furthermore, it also considers model uncertainty and therefore, inference is based on the proper uncertainty measures. However, I obtain further evidence that allows me to conclude that some variables such as population or life expectancy, in spite of having a statistically insignificant effect on growth, should be included as controls in growth regressions. This is so because the models that include these variables are the best models in fitting the data.

The remainder of the paper is organized as follows. Section 2 describes the construction of the likelihood function in the context of a partial adjustment with feedback panel data model. Monte Carlo evidence on the finite sample behavior of the estimator is provided in Section 3. In Section 4 I estimate some different specifications of empirical growth models with the proposed estimator. The obtained results when combining the estimator and model averaging techniques are presented in Section 5. Finally, Section 6 concludes.

2 Dynamic Panel Data with Feedback: Likelihood-Based Estimation

I consider the following panel data model:

$$y_{it} = \alpha y_{it-1} + x_{it}'\beta + w_i'\delta + \eta_i + \zeta_t + v_{it} \tag{1}$$

$$E\left(v_{it} \mid y_i^{t-1}, x_i^t, w_i, \eta_i\right) = 0 \qquad (t = 1, ..., T)(i = 1, ..., N) \tag{2}$$

where:

$$\begin{aligned} x_{it} &= \left(x_{it}^1, \ldots, x_{it}^k\right)' \\ w_i &= \left(w_i^1, \ldots, w_i^m\right)' \\ \beta &= (\beta_1, \ldots, \beta_k)' \\ \delta &= (\delta_1, \ldots, \delta_m)' \end{aligned}$$

The predetermined nature of the lagged dependent variable is considered in assumption (2). The model also relaxes the strict exogeneity assumption for the x variables that are also considered as predetermined (this is why we can refer to the model as having general predetermined variables). In particular, the assumption in (2) allows for feedback from lagged values of y to the current value for x. Moreover it implies lack of autocorrelation in v_{it} since lagged vs are linear combinations of the variables in the conditioning set.

I also include m strictly exogenous regressors that may or may not have temporal variation. In the remaining of the exposition I will assume that all the w variables have no variation within time. Despite allowing for time varying strictly exogenous w variables is straightforward in this context, in the spirit of Hausman and Taylor (1981) I prefer to stress the possibility of identifying the effect of time-invariant variables in addition to the unobservable time-invariant fixed effect. This is possible by assuming lack of correlation between the w variables and the unobservable fixed effects η_i.

Note that in addition to the individual specific fixed effects η_i, I have also included the term ζ_t in (1). That is to say, I am including time dummies in the model in order to capture unobserved common factors across units in the panel and therefore, I am not ruling out cross-sectional dependence. In the practice, this is done by simply working with cross-sectional de-meaned data. In the remaining of the exposition, I assume that all the variables are in deviations from their cross-sectional mean.

There are two possible parametrizations of the model given by (1)-(2). Under some distributional assumptions, both parametrizations give, for the same underlying model, two different log-likelihood functions with restrictions in the variance-covariance matrices. In this section I will present one of them, the denominated Simultaneous Equations Model (SEM) representation, since it will allow me to concentrate the free parameters of the resulting log-likelihood in order to make feasible its maximization. For additional insights, the other possible parametrization of the model, labeled as Full Covariance Structure (FCS) representation, can be found in the Appendix A.1.

In order to introduce the SEM parametrization, note that we may rewrite the model in (1)-(2) as a system of equations. In particular, we would have T structural form equations, one for each period, given by (1). Moreover, one could complete these T structural form equations with additional reduced form equations for the predetermined regressors. In both parametrizations we have a set of structural form equations and a set of reduced form equations. Hence, we can interpret this likelihood-based estimator as a sub-system LIML estimator. More concretely, the underlying idea of the SEM parametrization is to consider y_{i0} and x_{i1} as strictly exogenous variables in the system and, therefore, the feedback process will be captured by the variance-covariance matrix of the model. This is a very convenient representation because it will allow me to reduce the dimension of the problem by concentrating the log-likelihood of the system with respect to some reduced form parameters (Note that the number of parameters to be estimated is huge because I parametrize the whole feedback process and therefore, it is necessary to get rid of some of them in order to make feasible the estimation).

In the SEM parametrization y_{i0} and x_{i1} are considered as strictly exogenous variables, hence, the first equation of the system is:

$$\eta_i = \gamma_0 y_{i0} + x'_{i1}\gamma_1 + \epsilon_i \tag{3}$$

Note that in (3) we are implicitly assuming that $Cov(\eta_i, w_i) = 0$ in order to ensure identification of δ. Moreover, by substituing (3) in (1) the whole model can be writen as follows:

$$y_{i1} = (\alpha + \gamma_0)y_{i0} + x'_{i1}(\beta + \gamma_1) + w'_i\delta + \epsilon_i + v_{i1} \tag{4a}$$

and for $t = 2, ..., T$:

$$y_{it} = \alpha y_{i,t-1} + x'_{it}\beta + \gamma_0 y_{i0} + x'_{i1}\gamma_1 + w'_i\delta + \epsilon_i + v_{it} \tag{4b}$$

$$x_{it} = \pi_{t0}y_{i0} + \pi_{t1}x_{i1} + \pi^w_t w_i + \xi_{it} \tag{4c}$$

where ξ_{it}, γ_1 and π_{t0} are $k \times 1$ vectors, π_{t1} is a $k \times k$ matrix and π^w_t a $k \times m$ matrix.

In order to rewrite the system in matrix form, I define the following $T + (T-1)k \times 1$ vectors of data and errors for individual i:

$$\begin{aligned} R_i &= (y_{i1}, y_{i2}, \dots, y_{iT}, x'_{i2}, x'_{i3}, \dots, x'_{iT})' \\ U_i &= (\epsilon_i + v_{i1}, \dots, \epsilon_i + v_{iT}, \xi'_{i2}, \dots, \xi'_{iT})' \end{aligned}$$

Therefore I am now able to rewrite the model in matrix form as follows:

$$BR_i = \Pi z_i + U_i \tag{5}$$

where B and Π are matrices of coefficients defined below and z_i is the $(1 + k + m) \times 1$ vector of exogenous variables:

$$z_i = (y_{i0}, x'_{i1}, w'_i)'$$

Moreover, if I additionally define the following vectors:

$$\begin{aligned} R_{i1} &= (y_{i1}, y_{i2}, \dots, y_{iT})' \\ R_{i2} &= (x'_{i2}, x'_{i3}, \dots, x'_{iT})' \\ U_{i1} &= (\epsilon_i + v_{i1}, \dots, \epsilon_i + v_{iT})' \\ U_{i2} &= (\xi'_{i2}, \dots, \xi'_{iT})' \end{aligned}$$

it is then possible to rewrite:

$$\left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline 0 & I_{k-1} \end{array}\right)\left(\begin{array}{c} R_{i1} \\ \hline R_{i2} \end{array}\right) = \left(\begin{array}{c} \Pi_1 \\ \hline \Pi_2 \end{array}\right) z_i + \left(\begin{array}{c} U_{i1} \\ \hline U_{i2} \end{array}\right) \tag{6}$$

where:

$$B_{11} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\alpha & 1 & 0 & \dots & 0 \\ 0 & -\alpha & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -\alpha & 1 \end{pmatrix}_{T\times T} \qquad B_{12} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ -\beta' & 0 & \dots & 0 \\ 0 & -\beta' & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -\beta' \end{pmatrix}_{T\times k(T-1)}$$

$$\Pi_1 = \begin{pmatrix} \alpha+\gamma_0 & \beta'+\gamma_1' & \delta' \\ \gamma_0 & \gamma_1' & \delta' \\ \vdots & \vdots & \vdots \\ \gamma_0 & \gamma_1' & \delta' \end{pmatrix}_{T\times(1+k+m)} \qquad \Pi_2 = \begin{pmatrix} \pi_{20} & \pi_{21} & \pi_2^w \\ \vdots & \vdots & \vdots \\ \pi_{T0} & \pi_{T1} & \pi_T^w \end{pmatrix}_{k(T-1)\times(1+k+m)}$$

Considering this SEM parametrization, the variance-covariance matrix of the disturbance terms is as follows:

$$\Omega = Var(U_i) = Var\left(\begin{array}{c} U_{i1} \\ \hline U_{i2} \end{array}\right) = \left(\begin{array}{c|c} \Omega_{11} & \Omega_{12} \\ \hline \Omega_{21} & \Omega_{22} \end{array}\right) \tag{7}$$

The term Ω_{11} has the classical error-component form but allowing for time-series heteroskedasticity:

$$\Omega_{11} = \sigma_\epsilon^2 \iota\iota' + \begin{pmatrix} \sigma_{v_1}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{v_T}^2 \end{pmatrix}$$

where ι is a $T \times 1$ vector of ones.

On the other hand, Ω_{22} is the $(T-1)k \times (T-1)k$ covariance matrix that gathers all the contemporaneous and dynamic relationships between the x variables:

$$\Omega_{22} = \begin{pmatrix} \Sigma_{2,2} & & & \\ \Sigma_{2,3} & \Sigma_{3,3} & & \\ \vdots & \vdots & \ddots & \\ \Sigma_{2,T} & \Sigma_{3,T} & \dots & \Sigma_{T,T} \end{pmatrix}$$

where $\Sigma_{f,g}$ is the $k \times k$ covariance matrix between x_{if} and x_{ig}.

Finally, as previously stated, with the SEM parametrization the feedback process is captured through restrictions in the term Ω_{12}. In particular, given the assumptions above I can write:

$$cov(\epsilon_i, \xi_{it}) = \phi_t \qquad \forall t = 2, ..., T \tag{8a}$$

$$cov(v_{ih}, \xi_{it}) = \begin{cases} \psi_{h,t} & \text{if } h < t \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{8b}$$

where ϕ_t, $\psi_{h,t}$ and $\mathbf{0}$ are $k \times 1$ vectors. Therefore:

$$\Omega_{12} = \begin{pmatrix} \phi_2' + \psi_{1,2}' & \phi_3' + \psi_{1,3}' & \dots & \phi_T' + \psi_{1,T}' \\ \phi_2' & \phi_3' + \psi_{2,3}' & \dots & \phi_T' + \psi_{2,T}' \\ \phi_2' & \phi_3' & \dots & \phi_T' + \psi_{3,T}' \\ \vdots & \vdots & \ddots & \vdots \\ \phi_2' & \phi_3' & \dots & \phi_T' + \psi_{T-1,T}' \\ \phi_2' & \phi_3' & \dots & \phi_T' \end{pmatrix}_{T\times(T-1)k}$$

Under normal errors the log-likelihood for the model can be writen as[3]:

$$L \propto -\frac{N}{2}\ln\det(\Omega) - \frac{1}{2}tr\left(\Omega^{-1}U'U\right) \tag{9}$$

where U' is a $T + (T-1)k \times N$ matrix that consists of the U_i column vectors of each of the N individuals.

The number of parameters to be estimated in (9) is enormous. In order to make the problem feasible I will work with the concentrated log-likelihood with respect to the free parameters in the matrices Π_2 and Ω_{22}. See Appendix A.2 for more details on the concentration of the SEM log-likelihood.

3 Monte Carlo simulation

TBA

4 Empirical Growth Regressions

The neoclassical framework represents the basis for most empirical growth research. Departing from a generic one-sector growth model, in either its Solow-Swan or Ramsey-Cass-Koopmans variant, it is usual to assume that aggregate output obeys a Cobb-Douglas production function and then obtain a canonical cross-country growth regression of the form:

$$\gamma_i = \beta\ln y_{i0} + \psi X_i + \epsilon_i \tag{10}$$

where $\gamma_i = t^{-1}(\ln y_{it} - \ln y_{i0})$ represents the growth rate of output per worker between 0 and t. On the other hand, X_i is a vector of variables that represents not only the growth determinants suggested by the the Solow-Swan growth model but also additional determinants that allow for predictable heterogeneity in the steady state. These regressions are sometimes called Barro regressions, given Barro's extensive use of such regressions to study alternative growth determinants starting with Barro (1991). These kind of regressions

[3]Note that $\det(B) = 1$

have been widely used trying to address two major themes in the formal empirical analysis of growth: the identification of growth determinants and the question of convergence.

As previously stated, most of the growth econometrics literature is based on equation (10). The main objective of the present paper is to solve the problems that are still present in these empirical growth regressions from an econometric perspective. In particular, I address the issues of endogeneity, omitted variables, model uncertainty, measurement error, and, to some extent, parameter heterogeneity. By doing so, I will then be able to shed some light on the two issues mentioned above.

There is an important variant of the baseline empirical growth regression in (10) that can be called the canonical panel growth regression:

$$\ln y_{i,t} = (1+\beta)\ln y_{i,t-1} + \psi X_{i,t-1} + \eta_i + \zeta_t + v_{i,t} \qquad (i=1,...,N)(t=1,...,T) \tag{11}$$

where η_i is a country-specific fixed effect that allows considering unobservable heterogeneity across countries (since this term is country specific, we can interpret it as allowing for some kind of parameter heterogeneity across countries), and ζ_t is a period-specific shock common to all countries. The use of panel data in empirical growth regressions has many advantages with respect to cross-sectional regressions. First of all, the prospects for reliable generalizations in cross-country growth regressions are often constrained by the limited number of countries available, therefore, the use of within-country variation to multiply the number of observations is a natural response to this constraint. On the other hand, the use of panel data methods allows solving the inconsistency of empirical estimates which typically arises with omitted country specific effects which, if not uncorrelated with other regressors, lead to a misspecification of the underlying dynamic structure, or with endogenous variables which may be incorrectly treated as exogenous.

There are several issues to be treated in the panel growth regressions literature. Firstly, dependence of the lagged dependent variable and the regressors in $X_{i,t-1}$ with the country-specific fixed effect is allowed in virtually all previous panel studies. In this manner, the country-specific fixed effects are treated as parameters to be estimated and we condition on them, so, their distribution plays no role. This is the so-called fixed effects approach in contrast to the random effects approach that invokes a distribution for η and considers the effects independent of all the regressors in the model. Secondly, Knight et al. (1992) and Islam (1995) among others, have also consider the predetermined nature of the lagged dependent variable with respect to the transitory component of the error term $v_{i,t}$. This point refers to the fact that, by construction, all leads of $y_{i,t-1}$ are correlated with $v_{i,t}$ and, therefore, the within-groups estimator will produce biased estimates in the typical small-T growth panel. In particular, both studies employ the Π-matrix method of Chamberlain (1983). An important drawback of this method is that all the variables in the X vector are considered as strictly exogenous, *i.e.* all leads and lags of the variables are assumed to be uncorrelated with $v_{i,t}$. This consideration rules out the possibility of feedback from lagged income (*i.e.* $\ln y$) to current growth determinants such as the rate of investment or the rate of population growth (*i.e.* the x variables), which seems to be reasonable in the growth context. Finally, Caselli et al. (1996) and Benhabib and Spiegel (2000) among others, take into consideration the predetermined nature[4] of the x variables allowing for the mentioned feedback process. In particular, in order to estimate the model, they use generalized method of moments (GMM) following techniques advanced by Holtz-Eakin et al. (1988) and Arellano and Bond (1991). The assumptions that the error term is serially uncorrelated and that the explanatory variables are predetermined imply a set of moment restrictions that can be used in the context of GMM to generate consistent and efficient estimates of the parameters of interest. More concretely, the employed moment restrictions can be interpreted as an instrumental variables model where lagged levels of the variables are used as instruments for their first-differences. As Blundell and Bond (1998) pointed out, with persistent series such as GDP, lagged levels may be only weak instruments for the equation in first-differences. Thus, in spite of being consistent as N goes to infinity, this estimator is poorly behaved in finite samples. For this reason, these GMM estimates are not very reliable and have not received too much credit in the empirical growth literature. In order to solve this weak instruments problem, Bond et al. (2001) proposed, in the context of growth regressions, the use of the so-called system-GMM estimator introduced by Arellano and Bover (1995). However, this estimator requires the additional assumption of mean stationarity of the variables. Additional stationarity assumptions for solving this weak instruments problem are considered an *ad hoc* solution and not very appealing. In the growth regressions framework, this assumption is specially not desiderable since it may be interpreted as assuming that all the countries are in their steady state.

To the best of my knowledge there is no better alternative to estimate empirical panel growth regressions. The sub-system LIML estimator presented in the previous section is a good candidate for solving the problems described above. First of all, it considers the presence of country-specific fixed effects that may be correlated with both lagged income and growth determinants. Secondly, it also takes into consideration the predetermined nature not only of the lagged dependent variable but also of the growth determinants (*i.e.* feedback from lagged income to current growth determinants is allowed). Thirdly, as it is well-known, LIML estimators are free from finite sample biases caused by weak instruments. Moreover, measurement error considerations can be easily accommodated through additional restrictions on the variance-covariance matrix. On the other hand, it is important to remark that model uncertainty will be considered in the next section.

Given the above, the model to be estimated is given by the following equation and assumption:

$$y_{i,t} = \alpha y_{i,t-1} + \psi x_{i,t-1} + \eta_i + \zeta_t + v_{i,t} \tag{12a}$$

$$E\left(v_{i,t} \mid y_i^{t-1}, x_i^{t-1}, \eta_i\right) = 0 \qquad (i = 1, ..., N)(t = 1, ..., T) \tag{12b}$$

where $\alpha = 1 + \beta$, $y_{i,t}$ is the GDP per capita for country i in period t, $x_{i,t-1}$ is a $k \times 1$

[4]This predetermined nature is sometimes denominated weakly exogeneity in the growth literature.

vector of growth determinants, η_i is a country-specific fixed effect, ζ_t represents a set of time dummies and $v_{i,t}$ is the random disturbance term.

Given current data availability, it is now possible to use 10-year periods in panel growth regressions. This is so because typical sources of "growth data" such as Penn World Tables, cover a broad range of countries over the period 1960 to 2000. By using 10-year periods I aim to avoid the effect of business-cycle fluctuations and, therefore, focus on the long-term growth process. However, I will also present some estimations using 5-year periods data with similar results.

4.1 Revisiting the Solow-Swan Model

The baseline empirical growth regression is given by the basic neoclassical growth model, developed by Solow (1956) and Swan (1956). In the empirical counterpart of this model, the vector $x_{i,t-1}$ in (12a) includes proxies for the population growth rate (n), the rate of technological progress (g), the rate of depreciation of physical capital (d), and the saving rate (s). In particular, in my regressions, output is measured by GDP per capita at constant 2000 international prices from Penn World Tables 6.2 (PWT62). The saving rate (s) is proxied by the ratio of real domestic investment to GDP from PWT62. Finally, following Mankiw et al. (1992) and Caselli et al. (1996) among others, I choose 0.05 as a reasonable assessment of the value of $g + d$. Appendix A.3 contains more details about the employed data.

I have applied different estimation methods to the Solow-Swan model in two different panel settings, five-year periods and ten-year periods data. The obtained results are presented in Table 1. The bulk of the empirical growth regressions literature in based on cross-country OLS regressions as presented in columns (1) and (5). The within-groups (WG) estimator is a slight variant where given the availability of a panel dataset, country dummies can be included in order to allow for the presence of unobserved heterogeneity (*i.e.* country-specific fixed effects). The results when employing both OLS and WG estimators are in line with previous literature. The problem is that, as previously stated, these estimates are based on the wrong assumptions and thus they are only biased estimates of the real effects. On the other hand, the similarity between WG and diff-GMM estimates is interpreted as an indication of the presence of a weak instruments problem. This has been previously documented in Bond et al. (2001). As a result, in spite of being based on reasonable assumptions, the diff-GMM estimates are not reliable because they suffer from finite sample biases.

The sub-system LIML estimation procedure presented in this paper is applied to the basic Solow-Swan model and the results are shown in columns (4) and (8) of Table 1. Inspection of these columns makes clear the importance of the finite sample biases existent in previous differenced GMM estimates of this model. In contrast to previous panel estimates

Table 1: Solow-Swan Model Estimation Results

	Five-year data				Ten-year data			
	OLS	WG	diff GMM	sub-system LIML	OLS	WG	diff GMM	sub-system LIML
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable is $\ln(y_{i,t})$							
$\ln(y_{i,t-1})$	0.963	0.843	0.830	1.012	0.927	0.718	0.717	1.007
	(0.007)	(0.025)	(0.050)	(0.040)	(0.014)	(0.050)	(0.112)	(0.064)
$\ln(s_{i,t-1})$	0.088	0.091	0.035	0.095	0.167	0.166	0.009	0.209
	(0.010)	(0.018)	(0.034)	(0.027)	(0.019)	(0.036)	(0.085)	(0.033)
$\ln(n_{i,t-1}+g+d)$	−0.204	−0.137	0.128	0.020	−0.441	−0.327	0.557	0.111
	(0.041)	(0.071)	(0.108)	(0.111)	(0.085)	(0.163)	(0.325)	(0.205)
Implied λ	0.007	0.034	0.037	−0.002	0.008	0.033	0.033	−0.001
	(0.001)	(0.006)	(0.012)	(0.008)	(0.002)	(0.007)	(0.016)	(0.006)
Observations	584	584	511	584	292	292	219	292
Countries	73	73	73	73	73	73	73	73

Notes: In all columns a set of time dummies is included in the regressions. Columns (1) and (5) refer to the OLS estimation without country-specific fixed effects and all regressors considered as exogenous. In columns (2) and (6) the within-group estimator is employed and therefore fixed effects are included. However all regressors are assumed to be strictly exogenous. Finally, columns (3)-(4) and (7)-(8) present different estimates of the Solow-Swan version of the model in (12a)-(12b), where both fixed effects and weakly exogeneity are considered. In particular, columns (3) and (7) refer to the differenced GMM estimation and columns (4) and (8) present the estimation results when using the sub-system LIML estimator presented in Section 2. Standard errors are in parenthesis.

of the rate of convergence using the Solow-Swan framework, I obtain here that the speed of convergence is either low or zero across the countries in the sample. This is true when considering both five-year and ten-year periods. In particular, the point estimate for the convergence rate[5] is roughly zero in both cases. However, the 95% confidence intervals imply convergence rates that vary from −1.8% to 1.3% in the case of five-year periods data and from −1.3% to 1.2% in the case of ten-year data. This result indicates that previous panel studies such as Caselli et al. (1996), where the estimated rate of convergence was surprisingly high, were driven by finite sample biases. This conclusion will be reinforced in the remaining of the paper when Barro regressions and model uncertainty will be also taken into account.

By the same token, some differences also arise with respect to other parameter estimates. More concretely, while the estimate for $\ln(n_{i,t-1} + g + d)$ is similar in both diff-GMM and sub-system LIML (*i.e.* not significantly different from zero), the savings rate coefficient is positive, larger and significant in the case of sub-system LIML but insignificant when using diff-GMM. Moreover, its effect is always larger in the case of ten-year periods data.

[5] The convergence rate λ is obtained as follows: $\lambda = \frac{\ln \alpha}{-\tau}$ where τ is either 5 or 10. On the other hand, its standard error is calculated by the delta method.

4.2 Barro Regressions

Since Barro (1991), most of empirical growth regressions are based on a wide variety of specifications given by different variables included in the vector $x_{i,t-1}$ in (12a). In this subsection I will apply the sub-system LIML estimator together with OLS, WG and diff-GMM to two distinct panel cross-country growth regressions *la* Barro. In particular, I focus on the baseline specification of Barro and Lee (1994) as well as an alternative specification explained below.

The basic empirical framework of Barro regressions with panel data is given by equation (12a). Two kind of variables are included in theses regressions, first, initial levels of state variables measured at the beginning of the period (I will now focus on ten-year periods); and second, control or environmental variables, some of which are chosen by governments or private agents. For the baseline specification, as in Barro and Lee (1994), among the state variables I include the initial level of per capita GDP, the average number of years of secondary education, and the logarithm of life expectancy. The first is used to proxy the initial stock of physical capital, while the others are proxies for the initial level of human capital in the forms of educational attainment and health. Among the control variables, I include the domestic investment ratio (I/GDP) and the ratio of government consumption to GDP (G/GDP) as in Barro and Lee (1994). Given data availability in my sample period, the other two control variables are slightly different from those employed in the original specification but they capture similar effects. I consider the price of investment as a measure market prices distortions that exists in the economy and a polity composite index as a proxy of political freedom and stability. GDP, investment share, government consumption, and investment price are taken from PWT62. Secondary education is from Barro and Lee (2000), life expectancy from World Development Indicators 2005 and the polity index from the Polity IV project[6]. In the next section I will explain more about these and other state and control variables.

Table 2 shows the results. Columns (1)-(4) refer to the baseline specificacion previously described. In line with Solow-Swan estimation results, the main conclusion from these columns is that the rate of convergence is either very low or zero according to the sub-system LIML estimates. The 95% sub-system LIML confidence interval goes from -1.0% to 1.4%. On the other hand, the conclusions with respect to other explanatory variables may change a lot depending on the estimation method. For instance, government consumption has a negative and significative effect on growth according to the sub-system LIML estimates but not according to diff-GMM that suffer from finite sample bias.

In columns (5)-(8) I present the results from an alternative specification. This specification can be interpreted as follows: in the framework of the augmented Solow model (Mankiw et al. (1992)), the effect of political freedom and stability is tested. For this purpose, I add the polity index as an additional explanatory variable in a regression with the

[6] A more detailed description of the data sources and variables is in Appendix A.3

Table 2: Barro Regressions Estimation Results

	Baseline Specification Ten-year data				Alternative Specification Ten-year data			
	OLS	WG	diff GMM	sub-system LIML	OLS	WG	diff GMM	sub-system LIML
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable is $\ln(y_t)$							
$\ln(y_{t-1})$	0.845	0.683	0.842	0.977	0.937	0.679	0.709	1.049
	(0.021)	(0.052)	(0.075)	(0.060)	(0.018)	(0.052)	(0.100)	(0.101)
Education	0.040	0.039	0.055	0.030	0.013	0.011	0.008	−0.045
	(0.015)	(0.036)	(0.081)	(0.019)	(0.017)	(0.034)	(0.053)	(0.037)
ln (life expect)	0.829	0.478	0.709	0.862				
	(0.108)	(0.224)	(0.488)	(0.155)				
I/GDP	0.588	0.781	0.857	1.114	0.927	0.899	0.649	1.367
	(0.133)	(0.213)	(0.279)	(0.285)	(0.130)	(0.204)	(0.302)	(0.308)
G/GDP	−0.246	−0.465	−0.314	−0.546				
	(0.115)	(0.284)	(0.534)	(0.149)				
Inv. Price	−0.0004	−0.0007	−0.0008	−0.0010				
	(0.0002)	(0.0003)	(0.0006)	(0.0004)				
Polity	−0.042	−0.201	−0.260	−0.256	0.014	−0.185	−0.296	−0.201
	(0.041)	(0.061)	(0.083)	(0.067)	(0.043)	(0.061)	(0.087)	(0.226)
Pop. Growth					−5.973	−3.082	4.707	−0.775
					(1.388)	(2.295)	(3.504)	(5.669)
Implied λ	0.017	0.038	0.017	0.002	0.007	0.039	0.034	−0.005
	(0.003)	(0.008)	(0.009)	(0.006)	(0.002)	(0.008)	(0.014)	(0.010)
Observations	292	292	219	292	292	292	219	292
Countries	73	73	73	73	73	73	73	73

Notes: The baseline specification is the same as in Barro and Lee (1994) and the alternative specification is explained in the main text. In all columns a set of time dummies is included in the regressions. Columns (1) and (5) refer to the OLS estimation without country-specific fixed effects and all regressors considered as exogenous. In columns (2) and (6) the within-group estimator is employed and therefore fixed effects are included. However all regressors are assumed to be strictly exogenous. Finally, columns (3)-(4) and (7)-(8) present different estimates of two versions of the model in (12a)-(12b) where both fixed effects and weakly exogeneity are considered. In particular, columns (3) and (7) refer to the differenced GMM estimation and columns (4) and (8) present the estimation results when using the sub-system LIML estimator presented in Section 2. Standard errors are in parenthesis.

covariates suggested by the augmented version of the Solow-Swan model (*i.e.* initial GDP, a measure of human capital, domestic investment and the rate of population growth). In this manner, we can think of this specification as a Barro regression. Given this specification, the sub-system LIML 95% confidence interval for the convergence rate estimate goes from −2.4% to 1.4%. On the other hand, there are now some results that are different depending not only on the estimation method but also on the specification. For example, in the baseline specification, the effect of the polity index is estimated to be negative and significative while in the alternative specification is not significantly different from zero according to the sub-system LIML estimates.

Given the above, it is easy to imagine thousands of Barro regressions in which the convergence parameter estimate will be different across specifications and in which the effects of the explanatory variables will also be different. This would lead us to misleading conclusions even if we consider unbiased and consistent estimates for a given model because we do not know whether this is the true empirical model or not. This fact illustrates the need to take into consideration model uncertainty in empirical growth regressions. In the next section, I combine the sub-system LIML unbiased and consistent estimates for a given specification with model averaging techniques in order to address model uncertainty.

5 Model Uncertainty

Since omitted variables, endogeneity, parameter heterogeneity and measurement error issues have been addressed in the previous section, I now turn to the issue of model uncertainty. Model uncertainty arises because the lack of clear theoretical guidance on the choice of growth regressors results in a wide set of possible specifications. Therefore, researcher's uncertainty about the value of the parameter of interest in a growth regression exists at distinct two levels. The first one is the uncertainty associated with the parameter conditional on a given empirical growth model. This level of uncertainty is of course assessed in virtually every empirical study. What is not fully assessed is the uncertainty associated with the specification of the empirical growth model. It is typical for a given paper that the specification of the growth regression is taken as essentially known; while some variations of a baseline model are often reported, via different choices of control variables, standard empirical practice does not systematically account for the sensivity of claims about the parameter of interest to model choice.

Many researchers consider that the most promising approach to account for model uncertainty is to employ model averaging techniques to construct parameter estimates that formally address the dependence of model-specific estimates on a given model. In this context, Sala-i-Martin et al. (2004) employ their Bayesian Averaging of Classical Estimates (BACE) to determine which growth regressors should be included in linear cross-country growth regressions. In a pure Bayesian spirit, Fernandez et al. (2001) apply the Bayesian Model Averaging approach with different priors but the same objective. Given that both papers are cross-sectional studies, Moral-Benito (2007) extends the approach to a panel data setting taking into account the presence of country-specific fixed effects and the endogeneity of the lagged dependent variable. However, there is no paper considering at the same time model uncertainty and the predetermined nature of all growth determinants.

This paper is the first one in doing so. More concretely, in this section, model averaging techniques are combined with the likelihood-based estimator previously introduced in order to simultaneously address the issues of endogeneity, omitted variable bias, parameter heterogeneity, measurement error and model uncertainty. Thus, we will be able to obtain unbiased and consistent estimates of what we can denominate causal effects in the growth context, that take into consideration the dependence of model-specific estimates on a given empirical growth model and, therefore, the uncertainty at the two different levels mentioned above.

5.1 Growth Determinants

As previously mentioned, the augmented Solow-Swan model can be taken as the baseline empirical growth model. It comprises four determinants of economic growth, initial income, rates of physical and human capital accumulation, and population growth. In addition to those four determinants, Durlauf et al. (2005)'s survey of the empirical growth literature identifies 43 distinct growth theories and 145 proposed regressors as proxies; each of these theories is found to be statistically significant in at least one study. The set of growth determinants considered in this paper is only a subset of that identified by Durlauf et al. (2005). This is so because of three main reasons: (i) Data availability in the panel data context for the postwar period 1960-2000 is smaller than in the cross-sectional case. (ii) Since number of models to be estimated increases exponentially with the number of regressors considered and it is necessary to resort to numerical optimization methods for each model estimation, the problem would be computationally intractable if we include too many candidates. (iii) Finally, as found by Ciccone and Jarocinski (2007), the fewer the potential growth determinants considered, the smaller the sensivity of the results. Therefore, for the purpose of robustness, I focus on the subset of available growth determinants given by those variables that are more relevant from a policy maker perspective. This excludes from the analysis geographic variables such as the fraction of land area in geographical tropics, that in spite of being available, they are of little relevance from a policy perspective.

In particular, I consider here the following growth determinants[7]:

- Initial GDP: One of the main features of the neoclassical growth model is the prediction of a low (less than one) coefficient on initial GDP (*i.e.* it predicts conditional convergence). If the other explanatory variables are held constant, then the economy tends to approach (or not) its long-run position at the rate indicated by the magnitude of the coefficient.

- Investment Ratio: The ratio of investment to output represents the saving rate in the neoclassical growth model. In this model, a higher saving rate raises the steady-state level of output per effective worker and therefore increases the growth rate for a given starting value of GDP. Many empirical studies such as DeLong and Summers (1991)

[7]A more detailed description of the data and its sources can be found in Appendix A.3

have found an important positive effect of the investment ratio on economic growth.

- Education: In the neoclassical growth model, since the seminal work of Lucas (1988), the concept of capital is usually broadened from physical capital to include human capital. Education is the form of human capital that has generated most of the empirical work. In spite of the positive theoretical effect, many empirical studies have failed in finding such an effect. In particular I consider here the years of secondary education from Barro and Lee (2000).

- Life Expectancy: Another commonly considered form of human capital is health. In particular, the log of life expectancy at birth at the start of each period is typically used as an indicator of health status. There is a growing consensus that improving health can have a large positive impact on economic growth. For example, Gallup and Sachs (2001) argue that wiping out malaria in sub-Saharan Africa could increase per capita GDP growth by 2.6% a year.

- Population Growth: The steady-state level of output per effective worker in the neoclassical growth model is negatively affected by a higher rate of population growth because a portion of the investment is devoted to new workers rather than to raise capital per worker. However, this implication is not always confirmed when estimating empirical growth models.

- Investment Price: Since the seminal work of Agarwala (1983), it is often argued that distortions of market prices impact negatively on economic growth. Given the connection between investment and growth, such market interferences would be especially important if they apply to capital goods. Therefore, following Barro (1991) and Easterly (1993) among others, I consider the investment price level as a proxy for the level of distortions of market prices that exists in the economy.

- Trade Openness: The trade regime/external environment is captured by the degree of openness measured by the trade openness, imports plus exports as a share of GDP. It is often argued that a higher degree of trade openness increases the opportunity set of profitable investments and therefore promotes economic growth. Many authors such as Levine and Renelt (1992) and Frankel and Romer (1999) have considered this ratio.

- Government Consumption: Since the seminal work of Barro (1991), many authors have considered the ratio of government consumption to GDP as a measure of stability and distortions in the economy. The argument is that government consumption has no direct effect on private productivity but lower saving and growth through the distorting effects from taxation or government-expenditure programs.

- Polity Measure: The role of democracy in the process of economic growth has been the source of considerable research effort. However, there is no consensus about how the level of democracy in a country affects economic growth. Some researchers believe that an expansion of political rights (*i.e.* more democracy) fosters economic rights and tends thereby to stimulate growth. Others think that the growth-retarding aspects of democracy such as the heightened concern with social programs and income redistribution may be the dominant effect. Many authors such as Barro (1996) and Tavares and Wacziarg (2001) have empirically investigated this issue. In this paper I consider the Polity IV index of democracy/autocracy for analyzing the overall effect of democracy on growth.

- Population: Romer (1987, 1990) and Aghion and Howitt (1992) among others, developed theories of endogenous growth that imply some benefits from larger scale. In particular, if there are significant setup costs at the country level for inventing or adapting new products or production techniques, then the larger economies would, on this ground, perform better. This countrywide scale effect is tested by considering country's population in millions of people.

5.2 Bayesian Averaging of Maximum Likelihood Estimates (BAMLE)

The basic idea behind model averaging is to estimate the distribution of unknown parameters of interest across different models. The fundamental principle of Bayesian Model Averaging (BMA) is to treat models and related parameters as unobservable, and to estimate their distributions based on the observable data. In contrast to classical estimation, model averaging copes with model uncertainty by allowing for all possible models to be considered, which consequently reduces the biases of parameters and makes inference more reliable.

Formally, consider a generic representation of an empirical model of the form:

$$\Psi = \theta X + \epsilon \tag{13}$$

where Ψ is the dependent variable of interest, and X represents a set of covariates. Imagine that there exist potentially very many empirical models, each given by a different combination of explanatory variables (*i.e.* different vectors X), and each with some probability of being the 'true' model. Suppose we have K possible explanatory variables. We will have 2^K possible combinations of regressors, that is to say, 2^K different models - indexed by M_j for $j = 1, ..., 2^K$- which all seek to explain y -the data-.

In order to obtain parameter estimates that formally consider the dependence of model-specific estimates on a given model, BMA techniques construct point estimates that are a weighted average of all the 2^K model-specific estimates for a given parameter. The weights are given by the posterior probability of the model to be the 'true' model[8]. More concretely,

[8] A more detailed discussion of the BMA methodology can be found in Hoeting et al. (1999) and Koop (2003) among others.

I will construct a weighted average of maximum likelihood estimates as follows:

$$\widehat{\widehat{\theta}} = \sum_{j=1}^{2^K} P\left(M_j|y\right) \widehat{\theta}^j_{ML} \tag{14}$$

where $\widehat{\theta}^j_{ML}$ is the ML estimate for model j. In this particular case, the sub-system LIML estimator presented in Section 2. It is important to note at this point, that each of the models being considered here is comprised by a set of simultaneous equations. Therefore, the sub-system LIML estimator maximizes the joint density of all the $1 + 2^K$ variables for all the possible models conditional on the strictly exogenous variables (*i.e.* initial observations). Then, a regressor is excluded from a particular model by restricting to zero its coefficients in the structural form equation. By doing so, the densities of the different models are comparable.

As pointed out by Moral-Benito (2007), equation (14) is true if we either assume diffuse priors on the parameter space for any given sample size, or have a large sample for any given prior on the parameter space. Similarly, I also compute a weighted average of all the estimated variances across different models:

$$\widehat{\sigma}^2_1(\widehat{\widehat{\theta}}) = \sum_{j=1}^{2^K} P\left(M_j|y\right) \widehat{Var}\left(\widehat{\theta}^j_{ML}\right) \tag{15}$$

However, this measure does not take into consideration the variance of estimates between regressions. Therefore, in order to illustrate how this variance is considered, following Leamer (1978) I also compute:

$$\begin{aligned} \widehat{\sigma}^2_2(\widehat{\widehat{\theta}}) &= \sum\nolimits_{j=1}^{2^K} P\left(M_j|y\right) \widehat{Var}\left(\widehat{\theta}^j_{ML}\right) \\ &+ \sum\nolimits_{j=1}^{2^K} P\left(M_j|y\right) \left(\widehat{\theta}^j_{ML} - \widehat{\widehat{\theta}}\right)^2 \end{aligned} \tag{16}$$

Inspection of (16) shows that the variance incorporates both the estimated variances of the individual models as well as the variance in estimates of the θ's across different models. Hence, the uncertainty at the two different levels mentioned above is taken into account.

Moreover, the weights[9] (*i.e.* the posterior model probabilities $P\left(M_j|y\right)$) are based on the Schwarz asymptotic approximation to the Bayes Factor, and therefore:

$$P\left(M_j|y\right) = \frac{P\left(M_j\right)\left(NT\right)^{\frac{-k_j}{2}} f(y|\widehat{\theta}_j, M_j)}{\sum_{i=1}^{2^K} P\left(M_i\right)\left(NT\right)^{\frac{-k_i}{2}} f(y|\widehat{\theta}_i, M_i)} \tag{17}$$

where $f(y|\widehat{\theta}_j, M_j)$ is the maximized likelihood function for model j.

Table 3 presents the results when aplying the BAMLE methodology together with the sub-system LIML estimator. Therefore, both model uncertainty and endogeneity are taken into consideration.

[9]Unweighted counterparts of the three measures in equations (14)-(16) are not reported here but they are available upon request.

Table 3: BAMLE Results

	$\widehat{\overline{\theta}}$	$\widehat{\sigma}_1(\widehat{\overline{\theta}})$	$\widehat{\sigma}_2(\widehat{\overline{\theta}})$	% signif. and pos.	% signif. and neg.	Post. Incl. Probability
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable is $\ln(y_t)$					
$\ln(y_{t-1})$	0.930	0.064	0.094	99.6%	0.0%	-
I/GDP	0.949	0.302	0.356	96.1%	0.0%	63.4%
Education	0.033	0.019	0.039	58.2%	20.3%	56.1%
Pop. Growth	−0.566	1.227	2.451	34.4%	48.4%	55.3%
Population	0.0006	0.0007	0.0010	18.0%	2.7%	98.0%
Inv. Price	−0.0005	0.0004	0.0006	0.0%	35.9%	47.9%
Trade Openness	0.038	0.036	0.057	73.0%	10.5%	60.7%
G/GDP	0.048	0.073	0.143	47.3%	37.1%	60.3%
ln(life expect)	0.078	0.081	0.209	67.2%	18.0%	75.7%
Polity	−0.125	0.076	0.134	22.3%	59.8%	50.4%

Notes: In this table, the sub-system LIML estimator introduced in Section 2 is combined with the BAMLE methodology described in the main text. The sample covers the period 1960 to 2000 divided in 10-years subperiods. Column (1) reports the weighted average of the sub-system LIML estimates across all the possible models (*i.e.* it corresponds to equation (14)). Column (2) refers to the square root of the weighted average of all the estimated variances presented in equation (15). Column (3) presents the square root of the posterior variance calculated as in equation (16) and, therefore, taking into account the variance of the estimates across different models. In columns (4) and (5) I report the percentage of models in which the estimated coefficient is significative at the 5% level and positive or negative respectively. Finally, column (6) refers to the posterior inclusion probability of a variable to be included in the 'true' empirical growth model. It is calculated as the sum of the posterior model probabilities of all the models containing that variable. Finally, while the results on the table are based on the assumption of a prior expected model size equal to four, results with different prior expected model sizes are very similar and available upon request.

Regarding the issue of convergence, the point estimate of the rate of convergence of an economy to its steady state is 0.73%. This estimate is a weighted average of unbiased and consistent estimates across all possible empirical growth models. However, considering both levels of uncertainty described above (*i.e.* applying the delta method to the standard error in column (3)), the estimate of the rate of convergence is not significantly different from zero. Therefore I can not reject the null hypothesis of no conditional convergence across the countries in my sample[10]. This result casts doubt on the conventional wisdom of conditional convergence as a strong empirical regularity in the country level data. For example, early versions of endogenous growth theories (*e.g.* Romer (1987, 1990) and Aghion and Howitt (1992)) were criticized because in contrast to the neoclassical growth model, they no longer predicted conditional convergence.

[10]This result was previously found in Moral-Benito (2007), where model uncertainty and the endogeneity of the lagged dependent variable were considered.

On the other hand, one important conclusion from Table 3 is that model uncertainty matters. If one does not consider the uncertainty associated with the specification of the empirical growth model, then, even if we take into account all the possible empirical growth models, we may obtain misleading conclusions. For example, according to the average standard errors in column (2), one could conclude that there are some variables such as the polity indicator and the investment ratio, that have a significative 'causal' effect on growth. However, once the variance of the parameters across different models is considered (*i.e.* looking at the standard error in column (3)), the investment ratio is the only variable that has a significative effect on growth.

The empirical evidence on growth determinants seems to be conclusive for only one variable, the investment ratio. While the associated standard errors are not distributed according to the usual t-distribution, Sala-i-Martin et al. (2004) note that in most cases, having a ratio of posterior conditional mean to standard deviation (*i.e.* the ratio of $\widehat{\widehat{\theta}}$ to $\widehat{\sigma}_2(\widehat{\widehat{\theta}})$) around two in absolute value indicates an approximate 95-percent Bayesian coverage region that excludes zero. This 'pseudo-t' statistic would indicate that in the case of the investment ratio, its positive effect on growth is significantly different from zero. Moreover, in the 94.5% of the estimated models its coefficient was estimated to be positive and significant at the 95% level.

For the rest of the growth determinants the picture emerging from Table 3 is a bit pessimistic since little can be said about them once all the potential biases and inconsistencies have been addressed. Based on the mentioned 'pseudo-t' statistic, there is no variable with an estimated causal effect significantly different from zero. At this point it is important to remark the difference between correlations and causal effects. While previous BMA studies applied to growth regressions obtain correlations, I claim to obtain here estimates of what can be labeled as causal effects. This would mean that given the available data, despite of the existence of variables robustly correlated with growth (see for example Sala-i-Martin et al. (2004), Fernandez et al. (2001) and Moral-Benito (2007)), besides the investment ratio, little can be said about which variables cause economic growth once inference is based on the proper measures of uncertainty.

Finally, the posterior inclusion probability (PIP), the sum of the model probabilities of all the models containing a particular variable, is quite high for some variables. For instance, population, which captures scale effects, has a PIP of 98.0%. Therefore, in spite of being not significant, population should be included in empirical growth regressions as a control variable since the models including population are those with the highest probability of being the true empirical growth model. Other variables with high PIP that should be included are life expectancy, the investment ratio, a measure of trade openness and the government consumption ratio.

6 Concluding Remarks

Previous empirical growth studies may be subject to criticism because of econometric inconsistencies. On the one hand, cross-sectional studies suffer from omitted variables biases. In these studies, unobservable country-specific characteristics are omitted from the empirical model. These characteristics do not seem to be uncorrelated with other growth regressors, and their omission lead to a misspecification of the underlying dynamic structure. Panel data methods solve this problem by including country-specific fixed effects in the empirical model. Despite panel studies address the issue of omitted variables biases, they typically treat predetermined variables as strictly exogenous. The most promising approach to solve this problem was the use of differenced generalized method of moments applied to panel growth regressions. This method allows accommodating country-specific fixed effects and predetermined variables by means of moment conditions. However, because of a weak instruments problem in the growth context, these GMM estimates suffer from large finite sample biases. In this paper, I present a likelihood-based estimator in a panel data context that does not suffer from finite sample biases and considers both fixed effects and endogeneity issues.

On the other hand, the empirical growth literature faces the problem of model uncertainty. Theory does not offer conclusive guidance when selecting the correct empirical growth model. This causes an additional bias in the parameter estimates that must be treated as conditional on a specific model. Moreover, inference exercises are typically based on incorrect measures of uncertainty that do not take into consideration this problem. In this paper, I also combine the proposed estimator with model averaging techniques in order to address the issue of model uncertainty.

My results indicate that both model uncertainty and endogeneity matter in empirical growth regressions. This is so because the conclusions very much depend on whether you consider these issues or not. In particular, I claim that only after addressing both problems we can obtain reliable conclusions about two prominent questions in the empirical growth literature: what variables cause economic growth and, whether there exists conditional convergence or not.

Once model uncertainty and endogeneity issues are controled for, I conclude that the hypothesis of lack of conditional convergence can not be rejected (at least accross the countries in my sample (see Appendix A.3)). This result casts doubt on one of the main predictions of the neoclassical model of growth that has been traditionally accepted, the existence of convergence of the economies towards their steady states.

With regard to the causes of economic growth, according to my results, there is only one variable that robustly promotes growth, the investment ratio. This conclusion is based on unbiased and consistent estimates, and also on the correct measures of uncertainty for inference purposes. For the rest of growth determinants considered in this paper, the empirical evidence available is not enough to conclude whether they significantly cause growth or not. This would lead us to a pessimistic picture from a policy perspective since only through increasing investment we could surely promote economic growth. Nevertheless, an interesting line for future research could be to investigate which factors cause investment once all the potential econometric inconsistencies and biases are accounted for.

On the other hand, looking at the posterior inclusion probability of the variables, I conclude that some of them (*e.g.* population, life expectancy, the investment ratio, a measure of trade openness and the government consumption ratio) should always be included as controls in empirical growth regressions.

A Appendix

A.1 Full Covariance Structure (FCS) Representation

I present here the FCS parametrization of the model given by (1)-(2). It will consist on T structural equations as in the SEM parametrization. The main difference between both is that now, in the FCS representation, the whole feedback process is captured by the parameters in the reduced form equations instead of by restrictions in the variance-covariance matrix. The FCS parametrization is therefore as follows:

$$y_{i0} = w_i'\delta_y + c_y\eta_i + v_{i0} \tag{18a}$$

$$x_{i1} = \Delta_1 w_i + \gamma_{10}y_{i0} + c_1\eta_i + u_{i1} \tag{18b}$$

$$y_{i1} = \alpha y_{i0} + x_{i1}'\beta + w_i'\delta + \eta_i + v_{i1} \tag{18c}$$

and for $t = 2, ..., T$:

$$x_{it} = \Delta_t w_i + \gamma_{t0}y_{i0} + ... + \gamma_{t,t-1}y_{i,t-1} + \Lambda_{t1}x_{i1} + ... + \Lambda_{t,t-1}x_{i,t-1} + c_t\eta_i + u_{it} \tag{18d}$$

$$y_{it} = \alpha y_{i,t-1} + x_{it}'\beta + w_i'\delta + \eta_i + v_{it} \tag{18e}$$

Remark: Note that by writing the system as in (18a)-(18e) we are implicitly assuming that $Cov(\eta_i, w_i) = 0$, since otherwise I should have added the equation $\eta_i = w_i'\delta_\eta + e_i$ in order to complete the system. Therefore, assuming that $\delta_\eta = 0$ is enough to guarantee identification of δ in (1).

We have a system of $T(k+1)+1$ equations where δ_y is the $m \times 1$ vector $(\delta_y^1, \ldots, \delta_y^m)'$, c_y is a scalar and c_t and γ_{th} are the $k \times 1$ vectors:

$$c_t = (c_t^1, \ldots, c_t^k)' \qquad (t = 1, \ldots, T)$$

$$\gamma_{th} = (\gamma_{th}^1, \ldots, \gamma_{th}^k)' \qquad (t = 1, \ldots, T) \quad (h = 0, \ldots, T-1)$$

Moreover, Δ_t and Λ_{th} are the following $k \times m$ and $k \times k$ matrices:

$$\Delta_t = \begin{pmatrix} \delta_t^{11} & \ldots & \delta_t^{1m} \\ \vdots & \ddots & \vdots \\ \delta_t^{k1} & \ldots & \delta_t^{km} \end{pmatrix} \qquad \Lambda_{th} = \begin{pmatrix} \lambda_{th}^{11} & \ldots & \lambda_{th}^{1k} \\ \vdots & \ddots & \vdots \\ \lambda_{th}^{k1} & \ldots & \lambda_{th}^{kk} \end{pmatrix}$$

and u_{it} is a $k \times 1$ vector of prediction errors.

On the other hand, I also define the $T(k+1)+2$ column vector of errors:

$$\Xi_i^F = (\eta_i, v_{i0}, u_{i1}', v_{i1}, \ldots, u_{iT}', v_{iT})'$$

and the $T(k+1)+1 \times 1$ vector of data for individual i:

$$R_i^F = (y_{i0}, x_{i1}, y_{i1}, \ldots, x_{iT}, y_{iT})'$$

Finally, in order to rewrite the system in matrix form, we define the $T(k+1)+1 \times T(k+1)+1$ lower triangular matrix of coefficients B_F as:

$$B_F = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ -\gamma_{10} & I_k & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ -\alpha & -\beta' & 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ -\gamma_{20} & -\Lambda_{21} & -\gamma_{21} & I_k & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & -\alpha & -\beta' & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 & 0 \\ -\gamma_{T0} & -\Lambda_{T1} & -\gamma_{T1} & -\Lambda_{T2} & -\gamma_{T2} & \ldots & -\gamma_{T,T-1} & I_k & 0 \\ 0 & 0 & 0 & 0 & 0 & \ldots & -\alpha & -\beta' & 1 \end{pmatrix}$$

And the matrices D_F and C_F of orders $T(k+1)+1 \times T(k+1)+2$ and $T(k+1)+1 \times m$ respectively:

$$D_F = \begin{pmatrix} c_y & 1 & 0 & 0 & 0 & 0 & \ldots & 0 \\ c_1 & 0 & I_k & 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & \ldots & 0 \\ c_2 & 0 & 0 & 0 & I_k & 0 & \ldots & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_T & 0 & 0 & 0 & 0 & 0 & I_k & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad C_F = \begin{pmatrix} \delta_y' \\ \Delta_1 \\ \delta' \\ \vdots \\ \Delta_T \\ \delta' \end{pmatrix}$$

Given the above, I am now able to write the system in matrix form as follows:

$$B_F R_i^F = C_F w_i + D_F \Xi_i^F$$

where:

$$Var\left(\Xi_i^F\right) = \Omega_F = \begin{pmatrix} \sigma_\eta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{v_0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Sigma_{u_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{v_1} & 0 & 0 & 0 \\ & & & & \ddots & & \\ 0 & 0 & 0 & 0 & 0 & \Sigma_{u_T} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_{v_T} \end{pmatrix}_{T(k+1)+2 \times T(k+1)+2}$$

and Σ_{u_t} is a $k \times k$ matrix.

Therefore, the log-likelihood is:

$$\begin{aligned} L_F = & - \frac{N}{2} \ln \det \left(B_F^{-1} D_F \Omega_F D_F' B_F'^{-1} \right) \\ & - \frac{1}{2} tr \left\{ \left(B_F^{-1} D_F \Omega_F D_F' B_F'^{-1} \right)^{-1} \left[R^F - W (B_F^{-1} C_F)' \right]' \left[R^F - W (B_F^{-1} C_F)' \right] \right\} \end{aligned}$$

where R^F and X_t are the following matrices:

$$\begin{aligned} R^F &= \begin{pmatrix} Y_0 & X_1 & Y_1 & \ldots & X_T & Y_T \end{pmatrix}_{N \times T(k+1)+1} \\ X_t &= \left(X_t^1, \ldots, X_t^k \right)_{NXk} \end{aligned}$$

and W is the $N \times m$ matrix:

$$W = \begin{pmatrix} w_1' \\ w_2' \\ \vdots \\ w_N' \end{pmatrix}$$

A.2 Concentrated Likelihood using the SEM Parametrization

Maximizing the log-likelihood in (9) may be cumbersome (or even impossible depending on the number of available observations) since the dimension of the numerical optimization problem is enormous. In particular, the number of parameters to be estimated (p) in (9) is determined by the following expression:

$$p = 3 + 2k + T + (T-1)(2+k+m)k + \frac{(T-1)k[(T-1)k+1]}{2} + \sum_{r=1}^{T-1} rk$$

As an illustrative example, suppose we have a panel with $T = 5$, $k = 7$ and $m = 4$, then $p = 862$. This number is huge and may cause the problem to be intractable, but it can be drastically reduced by concentrating some free parameters of the model. In particular, for this illustrative example, the number of parameters after concentrating the log-likelihood is reduced from $p = 862$ to $p = 120$.

The log-likelihood function in (9) will be concentrated with relation to Ω_{22} and Π_2 under the assumption that both terms are unconstrained. The concentrated log-likelihood will then be maximized by means of numerical optimization with relation to B_{11}, B_{12}, Π_1, Ω_{11} and Ω_{12} that are all restricted.

By grouping the observations for all individuals in columns, the model can be writen as follows:

$$\left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline 0 & I_{k-1} \end{array}\right) \left(\begin{array}{c} R_1' \\ \hline R_2' \end{array}\right) = \left(\begin{array}{c} \Pi_1 \\ \hline \Pi_2 \end{array}\right) Z' + \left(\begin{array}{c} U_1' \\ \hline U_2' \end{array}\right)$$

First of all, we define:

$$\begin{aligned}
\Omega^{-1} &= \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}^{-1} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \\
F_{12} &= G_{12} G_{22}^{-1} \\
F_{21} &= F_{12}'
\end{aligned}$$

and then rewrite:

$$\begin{aligned}
\det \Omega &= \det \Omega_{11} / \det G_{22} \\
tr(\Omega^{-1} U' U) &= tr(\Omega_{11}^{-1} U_1' U_1) + 2tr(G_{12} U_2' U_1) + tr(G_{22} U_2' U_2) + tr(G_{12} G_{22}^{-1} G_{21} U_1' U_1)
\end{aligned}$$

Therefore, (9) can be writen as follows:

$$\begin{aligned}
L &\propto -\frac{N}{2} \ln \det \Omega_{11} + \frac{N}{2} \ln \det G_{22} - \frac{1}{2} tr(\Omega_{11}^{-1} U_1' U_1) - tr(F_{12} G_{22} U_2' U_1) \qquad (19) \\
&- \frac{1}{2} tr(G_{22} U_2' U_2) - \frac{1}{2} tr(F_{12} G_{22} F_{21} U_1' U_1)
\end{aligned}$$

Note that we can also write $\Omega_{11}^{-1} = G_{11} - G_{12} G_{22}^{-1} G_{21}$ and I have added and substracted the term $tr(G_{12} G_{22}^{-1} G_{21} U_1' U_1)$.

Step 1: Concentrating out Π_2

Noting that $U_2' = R_2' - \Pi_2 Z'$, we can maximize the likelihood in (19) with respect to Π_2 and obtain its ML estimate:

$$\widehat{\Pi}_2 = R_2' Z(Z'Z)^{-1} + F_{21} U_1' Z(Z'Z)^{-1}$$

Given $\widehat{\Pi}_2$ we can write:

$$\begin{aligned} \widehat{U}_2' U_1 &= R_2' Q U_1 - F_{21} U_1' M U_1 \\ \widehat{U}_2' \widehat{U}_2 &= R_2' Q R_2 + F_{21} U_1' M U_1 F_{12} \end{aligned}$$

where M is the projection matrix on the exogenous variables of the system and Q the annihilator:

$$\begin{aligned} M &= Z(Z'Z)^{-1} Z' \\ Q &= I_N - M \end{aligned}$$

Replacing in (19), we obtain L_2, the log-likelihood concentrated with respect to Π_2:

$$\begin{aligned} L_2 \quad \propto \quad & -\frac{N}{2} \ln \det \Omega_{11} + \frac{N}{2} \ln \det G_{22} - \frac{1}{2} tr(\Omega_{11}^{-1} U_1' U_1) \\ - \quad & \frac{1}{2} tr\{(R_2 + U_1 F_{12})' Q (R_2 + U_1 F_{12}) G_{22}\} \end{aligned} \tag{20}$$

Step 2: Concentrating out Ω_{22}

I now turn to the concentration of L_2 with relation to Ω_{22}. Note that the log-likelohood is now writen in terms of G_{22} and therefore, in practice I will obtain the concentrated likelihood with respect to G_{22} instead of Ω_{22}. However, since they are unconstrained, this is simply a matter of notation.

First, we define:

$$H = (R_2 + U_1 F_{12})' Q (R_2 + U_1 F_{12})$$

Therefore:

$$L_2 \quad \propto \quad -\frac{N}{2} \ln \det \Omega_{11} + \frac{N}{2} \ln \det G_{22} - \frac{1}{2} tr(\Omega_{11}^{-1} U_1' U_1) - \frac{1}{2} tr\{H G_{22}\}$$

By differentiating the log-likelihood function, we obtain:

$$\begin{aligned} \mathbf{d} L_2 &= \frac{N}{2} tr(G_{22}^{-1} \mathbf{d} G_{22}) - \frac{1}{2} tr(H \mathbf{d} G_{22}) \\ &= tr[(\frac{N}{2} G_{22}^{-1} - \frac{1}{2} H) \mathbf{d} G_{22}] = 0 \end{aligned}$$

This implies that:

$$\widehat{G}_{22}^{-1} = \frac{1}{N} H$$

and so the final concentrated log-likelihood is:

$$L_3 \propto -\frac{N}{2} \ln \det \Omega_{11} - \frac{1}{2} tr(\Omega_{11}^{-1} U_1' U_1) - \frac{N}{2} \ln \det(\frac{1}{N} H) \tag{21}$$

A.3 Data Appendix

Table 4: Variable Definitions and Sources

Variable	Source	Definition
GDP	PWT 6.2	Logarithm of GDP per capita (2000 US dollars at PP)
I/GDP	PWT 6.2	Ratio of real domestic investment to GDP
Education	Barro and Lee (2000)	Stock of years of secondary education in the total population
Pop. Growth	PWT 6.2	Average growth rate of population
Population	PWT 6.2	Population in millions of people
Inv. Price	PWT 6.2	Purchasing-power-parity numbers for investment goods
Trade Openness	PWT 6.2	Exports plus imports as a share of GDP
G/GDP	PWT 6.2	Ratio of government consumption to GDP
ln (life expect)	WDI 2005	Logarithm of the life expectancy at birth
Polity	Polity IV Project	Composite index given by the democracy score minus the autocracy score. Original range -10,-9,...,10, normalized 0-1.

Notes: All variables are available for all the countries in the sample (see table below) and for the whole period 1960-2000. PWT 6.2 refers to Penn World Tables 6.2 and it can be fount at http://pwt.econ.upenn.edu/. WDI 2005 refers to World Development Indicators 2005. Data from Barro and Lee (2000) is available at http://www.cid.harvard.edu/ciddata/ciddata.html. Finally, data from the Polity IV Project can be downloaded from http://www.systemicpeace.org/polity/polity4.htm.

Table 5: List of Countries

Algeria	Indonesia	Peru	Argentina
Iran	Philippines	Australia	Ireland
Portugal	Austria	Israel	Rwanda
Belgium	Italy	Senegal	Benin
Jamaica	Singapore	Bolivia	Japan
South Africa	Brazil	Jordan	Spain
Cameroon	Kenya	Sri Lanka	Canada
Lesotho	Sweden	Chile	Malawi
Switzerland	China	Malaysia	Syria
Colombia	Mali	Thailand	Costa Rica
Mauritius	Togo	Denmark	Mexico
Trinidad & Tobago	Dom. Republic	Mozambique	Turkey
Ecuador	Nepal	Uganda	El Salvador
Netherlands	United Kingdom	Finland	New Zealand
United States	France	Nicaragua	Uruguay
Ghana	Niger	Venezuela	Greece
Norway	Zambia	Guatemala	Pakistan
Zimbabwe	Honduras	Panama	India
Paraguay			

References

Agarwala, R. (1983): "Price Distortions and Growth in Developing Countries," *World Bank staff working papers*, No. 575.

Aghion, P. and P. Howitt (1992): "A Model of Endogenous Growth through Creative Destruction," *Econometrica*, 60, 323–351.

Arellano, M. and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277–297.

Arellano, M. and O. Bover (1995): "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29–52.

Barro, R. (1991): "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106, 407–443.

——— (1996): "Democracy and Growth," *Journal of Economic Growth*, 1, 1–27.

Barro, R. and J. W. Lee (1994): "Losers and Winners in Economic Growth," *Proceedings of the World Bank Annual Conference on Development Economics*, Washington D.C., 267–297.

——— (2000): "International Data on Educational Attainment: Updates and Implications," *CID Working Paper*, No. 042.

Benhabib, J. and M. M. Spiegel (2000): "The Role of Financial Development in Growth and Investment," *Journal of Economic Growth*, 5, 341–360.

Blundell, R. and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115–143.

——— (2000): "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19, 321–340.

Bond, S. R., A. Hoeffler, and J. Temple (2001): "GMM Estimation of Empirical Growth Models," *CEPR Discussion Papers*, 3048.

Caselli, F., G. Esquivel, and F. Lefort (1996): "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," *Journal of Economic Growth*, 1, 363–389.

Chamberlain, G. (1983): "Panel Data," in *Handbook of Econometrics*, ed. by Z. Griliches and M. Intriligator, Amsterdam: North-Holland, 1247–1318.

Ciccone, A. and M. Jarocinski (2007): "Determinants of Economic Growth: Will Data Tell?" *CEPR Discussion Paper No. 6544.*

DeLong, J. and L. Summers (1991): "Equipment Investment and Economic Growth," *Quarterly Journal of Economics*, 106, 445–502.

Durlauf, S. N., P. A. Johnson, and J. R. W. Temple (2005): "Growth Econometrics," in *Handbook of Economic Growth*, ed. by P. Aghion and S. Durlauf, Elsevier, vol. 1, chap. 8, 555–677.

Easterly, W. (1993): "How Much Do Distortions Affect Growth?" *Journal of Monetary Economics*, 32, 187–212.

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

Frankel, J. and D. Romer (1999): "Does Trade Cause Growth?" *American Economic Review*, 89, 379–399.

Gallup, J. and J. Sachs (2001): "The Economic Burden of Malaria," *American Journal of Tropical Medicine and Hygiene*, 64, 85–96.

Hausman, J. A. and W. E. Taylor (1981): "Panel Data and Unobservable Individual Effects," *Econometrica*, 49, 1377–1398.

Hoeting, J., D. Madigan, A. Raftery, and V. C. (1999): "Bayesian Model Averaging: A Tutorial," *Statistical Science*, 14, 382–417.

Holtz-Eakin, D., W. Newey, and H. S. Rosen (1988): "Estimating vector autoregressions with panel data," *Econometrica*, 56, 1371–1395.

Islam, N. (1995): "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 1127–1170.

Knight, M., N. Loayza, and D. Villanueva (1992): "Testing the Neoclassical Theory of Economic Growth: A Panel Data Approach," *IMF Staff Papers*, 40, 512–541.

Koop, G. (2003): *Bayesian Econometrics*, Wiley-Interscience.

Leamer, E. (1978): *Specification Searches*, New York: John Wiley & Sons.

Levine, R. and D. Renelt (1992): "A sensivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82, 942–963.

Lucas, R. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3–42.

Mankiw, N. G., D. Romer, and D. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407–437.

Moral-Benito, E. (2007): "Determinants of Economic Growth: A Bayesian Panel Data Approach," *CEMFI Working Paper Series No. 0719.*

Romer, P. (1987): "Growth Based on Increasing Returns due to Specialization," *American Economic Review*, 77, 56–62.

——— (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, 71–102.

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

Solow, R. (1956): "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 65–94.

Swan, T. (1956): "Economic Growth and Capital Accumulation," *Economic Record*, 32, 334–361.

Tavares, J. and R. Wacziarg (2001): "How Democracy Affects Growth," *European Economic Review*, 45, 1341–1378.