WPS4830

Policy Research Working Paper 4830

Determinants of Economic Growth

A Bayesian Panel Data Approach

Enrique Moral-Benito

The World Bank
Development Research Group
Macroeconomics and Growth Team
February 2009

Abstract

Model uncertainty hampers consensus on the key determinants of economic growth. Some recent cross-country, cross-sectional analyses have employed Bayesian Model Averaging to address the issue of model uncertainty. This paper extends that approach to panel data models with country-specific fixed effects. The empirical results show that the most robust growth determinants are the price of investment goods, distance to major world cities, and political rights. This suggests that growth-promoting policy strategies should aim to reduce taxes and distortions that raise the prices of investment goods; improve access to international markets; and promote democracy-enhancing institutional reforms. Moreover, the empirical results are robust to different prior assumptions on expected model size.

This paper—a product of the Growth and the Macroeconomics Team, Development Research Group—is part of a larger effort in the department to assess the determinants of economic growth. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at emoral@cemfi.es.

Determinants of Economic Growth: A Bayesian Panel Data Approach*

Enrique Moral-Benito†
CEMFI

*This paper was completed during my stay at the World Bank's research department. I would like to thank Luis Servén and Manuel Arellano for his overall guidance and insightful comments. I also thank Roberto León, Eduardo Ley and Ignacio Sueiro for their help and advice. All errors are my own.

†E-mail address: emoral@cemfi.es, or enrique.moral@gmail.com

1 Introduction

Over the last two decades, hundreds of empirical studies have attempted to identify the determinants of growth. This is not to say that growth theories are of no use for that purpose. Rather, the problem is that growth theories are, using a term due to Brock and Durlauf (2001), open-ended. This means that different growth theories are typically compatible with one another. For example, a theoretical view holding that trade openness matters for economic growth is not logically inconsistent with another theoretical view that emphasizes the role of geography in growth. This diversity of theoretical views makes it hard to identify the most effective growth-promoting policies. The aim of this paper is to shed some light on this issue.

From an empirical point of view, the problem this literature faces is known as model uncertainty, which emerges because theory does not provide enough guidance to select the proper empirical model. In the search for a satisfactory statistical model of growth, the main area of effort has been the selection of appropiate variables to include in linear growth regressions. The cross-country regression literature concerned with this task is enormous: a huge number of papers have claimed to have found one or more variables correlated with the growth rate, resulting in a total of more than 140 variables proposed as growth determinants.

A more specific issue was raised by Levine and Renelt (1992). From an extreme-bounds analysis, they concluded that very few variables were robustly correlated with growth. In contrast, Sala-i-Martin (1997) constructed weighted averages of OLS coefficients and found that some were fairly stable across specifications.

Many researchers consider that the most promising approach to accounting 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, Doppelhofer and Miller (2004) -henceforth SDM- employ their Bayesian Averaging of Classical Estimates (hereafter, BACE) to determine which growth regressors should be included in linear cross-country growth regressions, making an attempt to confirm in a Bayesian-inspired framework the results obtained by Sala-i-Martin (1997). In a pure Bayesian spirit, Fernandez, Ley and Steel (2001a) -henceforth FLS- apply the Bayesian Model Averaging approach with different priors but the same objective. Moreover, both methodologies allow constructing a ranking of variables ordered by their *robustness* as growth determinants. In spite of the focus on robustness of this approach, Ley and Steel (2007) show that the results are fairly sensitive to the use of different prior assumptions. Moreover, Ciccone and Jarocinski (2005) employ exactly the same methodologies and conclude that the list of growth determinants emerging from these approaches is sensitive to arguably small variations in the international income data used in the estimations.

The main objective of this paper is to extend the Bayesian Model Averaging

(BMA) methodology to a panel data framework. The use of panel data in empirical growth regressions has many advantages with respect to typical cross-country 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 to solve 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. Since the seminal work of Islam (1995), a lot of studies such as Caselli, Esquivel and Lefort (1996) have employed panel data models with country specific effects in empirical growth regressions.

In our case, to simultaneously address both omitted variable bias and issues of endogeneity, we employ a Maximum Likelihood estimator which is able to use the within variation across time and also the between variation across countries.

Against this background, the paper presents a novel approach, Bayesian Averaging of Maximum Likelihood Estimates (BAMLE), which is easy to interpret and easy to apply since it only requires the elicitation of one hyper-parameter, the expected model size, m. Moreover, the impact of different prior assumptions about m is minimal with the prior structure employed. Our methodology is similar to the BACE approach by SDM, but given the use of a maximum likelihood estimator, BAMLE is more flexible and it can be applied to a broader range of situations. In fact, under the assumption of spherical disturbances, BACE can be considered a particular case of BAMLE.

On the other hand, empirical results indicate that the sensivity of the list of robust growth determinants emerging from our approach to the choice of alternative sources of international income data is considerably smaller than found in the previous literature. The reason is that the number of potential regressors we include in our dataset is much smaller than the number considered in previous studies. Therefore, we conclude that the sensitivity of the results to variations in the source of international income data found by Ciccone and Jarocinski (2005) is also present when we consider country specific effects. However, given our results, we can also conclude that the fewer the regressors the smaller the sensitivity. For the purposes of *robustness*, this suggests that the set of candidate variables should avoid inclusion of multiple proxies for the same theoretical effect.

The remainder of the paper is organized as follows. Section 2 describes the BMA methodology and extends to the panel data case the prior structures proposed by SDM and FLS. Section 3 constructs the likelihood function, describes the use of the BIC approximation in the BMA context, and introduces the prior assumptions employed for implementation of the BAMLE approach. In Section 4 we briefly describe the data set. The empirical results employing two different sources for international income data (World Development Indicators 2005 -WDI 2005- and Penn World Table 6.2 -PWT 6.2-) are presented in Section 5. The final section concludes.

2 Bayesian Model Averaging

A generic representation of the canonical growth regression is:

$$\gamma = \theta X + \varepsilon, \tag{1}$$

where γ is the vector of growth rates, and X represents a set of growth determinants, including those originally suggested by Solow as well as others [1]. There exist potentially very many empirical growth models, each given by a different combination of explanatory variables, and each with some probability of being the 'true' model. This is the starting point of the Bayesian Model Averaging method.

However, there is one variable for which theory offers strong guidance, and is therefore exempt from the problem of model uncertainty: initial GDP, which should always be included in growth regressions (see Durlauf, Johnson and Temple 2005). As a result, in the remainder of the paper initial GDP will be included with probability 1 in all models under consideration.

Using the Bayesian jargon, a model is formally defined by a likelihood function and a prior density. 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-. M_j depends upon parameters θ^j. In cases where many models are being entertained, it is important to be explicit about which model is under consideration. Hence, the posterior for the parameters calculated using M_j is written as:

$$g\left(\theta^j|y, M_j\right) = \frac{f\left(y|\theta^j, M_j\right) g\left(\theta^j|M_j\right)}{f\left(y|M_j\right)}, \tag{2}$$

and the notation makes clear that we now have a posterior, a likelihood, and a prior for each model. The logic of Bayesian inference suggests that we use Bayes' rule to derive a probability statement about what we do not know (*i.e.* whether a model is correct or not) conditional on what we do know (*i.e.* the data). This means the posterior model probability can be used to assess the degree of support for M_j. Given the prior model probability $P(M_j)$ we can calculate the posterior model probability using Bayes Rule as:

$$P\left(M_j|y\right) = \frac{f\left(y|M_j\right) P\left(M_j\right)}{f\left(y\right)}. \tag{3}$$

Since $P(M_j)$ does not involve the data, it measures how likely we believe M_j to be the correct model before seeing the data. $f(y|M_j)$ is often called the

[1]The inclusion of additional control variables to the regression suggested by the Solow (or augmented Solow) model can be understood as allowing for predictable and additional heterogeneity in the steady state

marginal (or integrated) likelihood, and is calculated using (2) and a few simple manipulations. In particular, if we integrate both sides of (2) with respect to θ^j, use the fact that $\int g\left(\theta^j|y, M_j\right) d\theta^j = 1$ (since probability density functions integrate to one), and rearrange, we obtain:

$$f(y|M_j) = \int f\left(y|\theta^j, M_j\right) g\left(\theta^j|M_j\right) d\theta^j. \tag{4}$$

The quantity $f(y|M_j)$ given by equation (4) is the marginal probability of the data, because it is obtained by integrating the joint density of (y, θ^j) given y over θ^j. The ratio of integrated likelihoods of two different models is the Bayes Factor and it is closely related to the likelihood ratio statistic, in which the parameters θ^j are eliminated by maximization rather than by integration.

Moreover, considering θ a function of θ^j for each $j = 1, ..., 2^K$, (*i.e.* for each model j,θ is defined as the vector θ^j augmented with zeros for those regressors not included in model j) we can also calculate the posterior density of the parameters for all the models under consideration:

$$g(\theta|y) = \sum_{j=1}^{2^K} P(M_j|y)\, g(\theta|y, M_j) \tag{5}$$

If one is interested in point estimates of the parameters, one common procedure is to take expectations across (5):

$$E(\theta|y) = \sum_{j=1}^{2^K} P(M_j|y)\, E(\theta|y, M_j). \tag{6}$$

Following Leamer (1978), we calculate the posterior variance as:

$$\begin{aligned} V(\theta|y) &= \sum_{j=1}^{2^K} P(M_j|y)\, V(\theta|y, M_j) + \\ &\quad + \sum_{j=1}^{2^K} P(M_j|y) \left(E(\theta|y, M_j) - E(\theta|y)\right)^2. \end{aligned} \tag{7}$$

Inspection of (7) shows that the posterior variance incorporates both the estimated variances of the individual models as well as the variance in estimates of the θ's across different models.

In words, the logic of Bayesian inference implies that one should obtain results for every model under consideration and average them using appropiate weights. However, implementing Bayesian Model Averaging can be difficult since the number of models under consideration -2^K-, is often huge. This has led to various algorithms which do not require dealing with every possible model. In particular we will employ the so called Markov Chain Monte Carlo Model Composition (MC^3) algorithm (see the Computational Appendix for more details).

Given the above, we are now ready to introduce our measure of robustness. We estimate the posterior probability that a particular variable h is included in the regression, and we interpret it as the probability that the variable belongs in the true growth model. In other words, variables with high posterior probabilities of being included are considered as *robust* determinants of economic growth. This is called the *posterior inclusion probability* for variable h, and it is calculated as the sum of the posterior model probabilities for all of the models including that variable:

$$\text{posterior inclusion probability} = P(\theta_h \neq 0|y) = \sum\nolimits_{\theta_h \neq 0} P(M_j|y). \tag{8}$$

2.1 BACE-SDM Approach in a Panel Data Context

For a given group of regressors, that is, for a given model M_j, the estimated econometric model consists of the following equation and assumptions:

$$\begin{aligned} y_{it} - y_{it-\tau} &= \rho y_{it-\tau} + x_{it}'^{j}\beta^j + \eta_i + \zeta_t + v_{it} \quad (t = 1, ..., T)\ (i = 1, ..., N) \quad (9) \\ y_{it} &= \alpha y_{it-\tau} + x_{it}'^{j}\beta^j + \eta_i + \zeta_t + v_{it} \quad (\alpha = \rho + 1) \end{aligned}$$

$$E\left(v_i|y_i, x_i^j, \eta_i\right) = 0, \tag{A1}$$

where $v_i = (v_{i1}, ..., v_{iT})'$, $x_i^j = \left(x_{i1}^j, ..., x_{iT}^j\right)'$ and $y_i = (y_{i1}, ..., y_{iT})'$. We observe y_{it} (the log of per capita GDP for country i in period t) and the k^jx1 vector of explanatory variables x_{it}^j included in model M_j, but not η_i, which is an unobservable time-invariant regressor. Additionally, we assume:

$$Var\left(v_i|y_i, x_i^j, \eta_i\right) = \sigma^2 I_T. \tag{A2}$$

Under assumptions (A1) and (A2), the within-group estimator (henceforth, WG) is the optimal estimator of α and β^j for a given model.

Note that in addition to the individual specific fixed effect η_i, we have also included the term ζ_t in (9). That is to say, we are including time dummies in the model in order to capture unobserved common factors across countries and therefore we are not ruling out cross-sectional dependence. In the practice, this is done by simply working with cross-sectionally de-meaned data. In the remaining of the exposition, we assume that all the variables are in deviations from their cross-sectional mean.

Following Sala-i-Martin *et al.* (2004) we have implemented the denominated BACE approach in this context. The idea of BACE is to assume diffuse priors (as an indication of our ignorance) and make use of the result that, in the linear regression model, for a given model M_j, standard diffuse priors and Bayesian regression yield posterior distributions identical to the classical sampling distribution of OLS.

With the assumptions stated above we can rewrite (6) as:

$$E(\theta|y) = \sum\nolimits_{j=1}^{2^K} P(M_j|y)\,\widehat{\theta}^j, \tag{10}$$

where $\widehat{\theta}^j$ is the WG[2] estimate for θ with the regressor set that defines model j.

[2] Although assumption (A1) does not hold by definition in this context, we should remark that this is the easiest way of applying the methodology to panel data estimates and we can consider it as the starting point of our research.

Moreover, as the posterior odds' behavior is problematic with diffuse priors[3], SDM propose to use instead the Schwarz aymptotic approximation to the Bayes factor; therefore:

$$P\left(M_j|y\right) = \frac{P\left(M_j\right)\left(NT\right)^{-k^j/2} SSE_j^{-(NT)/2}}{\sum_{i=1}^{2^K} P\left(M_i\right)\left(NT\right)^{-k^i/2} SSE_i^{-(NT)/2}}, \tag{11}$$

where NT is the number of observations, K is the total number of regressors, k^j is the number of regressors included in model j and SSE_j is the sum of squared residuals of the j-model's regression. Regarding the prior model size (W), the BACE approach assumes that each variable is independently included in a model:

$$\begin{aligned} W &\sim Bin\left(K, \xi\right) \\ E\left(W\right) &= K\xi \Rightarrow \xi = \frac{m}{K}. \end{aligned} \tag{12}$$

Note that with this prior structure, the researcher only needs to fix the prior expected model size m which implies different prior inclusion probabilities for a given regressor (ξ).

2.2 BMA-FLS Approach in a Panel Data Context

One question that arises when we think in terms of Bayesian econometrics is how sensitive are the results to the choice of priors by the researcher? In this section, instead of the BACE approach based on diffusse priors, we implement the full Bayesian approach with the benchmark priors proposed by Fernández, Ley and Steel (2001b). These priors can be easily applied to the panel data case (fixed-effects model) if we rewrite the M_j model in the previous section as:

$$y_{it} = \alpha y_{it-\tau} + x_{it}^{\prime j}\beta^j + \phi_1 D_1 + ... + \phi_N D_N + \zeta_t + v_{it} \ (t = 1, ..., T) \ (i = 1, ..., N), \tag{13}$$

where the coefficients ($\phi_1 ... \phi_N$) are the individual unobservable effects for each country, ($D_1 ... D_N$) are N dummy regressors and again, all variables will be in deviations from their cross-sectional means given the presence of the time dummy ζ_t. Assumptions (A1) and (A2) also hold here, and the error term is supposed to follow a normal distribution. Fernández *et al.* (2001b) propose a natural conjugate prior distribution which allows employing the exact Bayes factor instead of using asymptotic approximations. For the variance parameter, which is common for all the models under consideration, the prior is improper and non-informative:

$$p\left(\sigma\right) \propto \sigma^{-1}. \tag{14}$$

[3] If we use noninformative priors for parameters not common to all the considered models, the posterior odds ratio will always lend overwhelming support for the model with fewer parameters, regardless of the data.

The g-prior (Zellner (1986)) for the slope parameters is a normal density with zero mean and covariance matrix equal to:

$$\sigma^2 \left(g_0 Z'^j Z^j\right)^{-1}, \tag{15}$$

where $Z^j = (y_{-1}, x^j, D_1, ..., D_N)$ and:

$$g_0 = \min\left(\frac{1}{NT}, \frac{1}{(k^j + N)^2}\right).$$

With this prior, both the posterior for each model and the Bayes factor have a closed form. Concretely, the Bayes factor (the ratio of integrated likelihoods) for model M_j versus model M_i is given by:

$$B_{ji} = \left(\frac{g_{oj}}{1+g_{oj}}\right)^{\frac{k_j+1}{2}} \left(\frac{g_{oi}+1}{g_{oi}}\right)^{\frac{k_i+1}{2}} \left(\frac{\frac{1}{g_{oi}+1} SSE_i + \frac{g_{oi}}{g_{oi}+1}(y'y)}{\frac{1}{g_{oj}+1} SSE_j + \frac{g_{oj}}{g_{oj}+1}(y'y)}\right)^{\frac{NT}{2}}. \tag{16}$$

Once we have specified the distribution of the observables given the parameters and the prior for these parameters, we only need to define the prior probabilities for each of the models. In particular, FLS assume that every model has the same *a priori* probability of being the true model:

$$P(M_j) = 2^{-K}. \tag{17}$$

The prior in (17) is the Binomial prior of SDM but employing $m = K/2$ instead of $m = 7$[4].

2.3 On the Effect of Prior Assumptions

We have presented and described two different prior structures employed in the BMA context. Both approaches give very similar results, and this is often misinterpreted as a symptom of robustness with respect to prior assumptions. Ley and Steel (2007) show that this similarity arises mostly by accident. The reason is that the different choices of the prior inclusion probability of each variable (ξ) – treated as fixed in both approaches – compensates the different penalties to larger models implied by the diffuse priors of SDM and the informative g-priors of FLS.

The effect of weakly-held prior views (as those that apply in the growth regression context) should be minimal. In search of this minimal effect, Ley and Steel (2007) propose a model prior specification and model size (W) given by the following assumptions:

$$W \sim Bin(K, \xi) \tag{18}$$

[4] This represents another difference with respect to the priors of the BACE-SDM approach in the previous subsection. Note that Sala-i-Martin et. al. (2004) propose $m = 7$ as a reasonable prior mean model size in the cross-country context. Here, we propose $m = 5$ for the panel data case.

$$\xi \sim Be\left(a, b\right), \tag{19}$$

where $a, b > 0$ are hyper-parameters to be fixed by the researcher. The difference with respect to SDM and FLS is to make ξ random rather than fixed. Model size W will then satisfy:

$$E\left(W\right) = \frac{a}{a+b}K. \tag{20}$$

The prior model size distribution generated in this way is the so-called Binomial-Beta distribution. Ley and Steel (2007) propose to fix $a = 1$ and $b = (K - m)/m$ through equation (20), so we only need to specify m, the prior mean model size, as in the BACE-SDM and BMA-FLS approaches.

As shown by Ley and Steel (2007), this prior specification with ξ random rather than fixed implies a substantial increase in prior uncertainty about model size, and makes the choice of m much less critical. Moreover, as we shall see later, with random ξ the effects of prior assumptions are much less severe.

3 Bayesian Averaging of Maximum Likelihood Estimates (BAMLE)

The BAMLE approach is based on averaging maximum likelihood estimates in a Bayesian spirit, *i.e.*, we rewrite equation (6) as follows:

$$E\left(\theta|y\right) = \textstyle\sum_{j=1}^{2^K} P\left(M_j|y\right)\widehat{\theta}^j_{ML}. \tag{21}$$

where $\widehat{\theta}^j_{ML}$ is the maximum likelihood estimate for θ in model j.

The argument behind equation (21) is twofold: *(i)* assuming diffuse priors on the parameter space of a given model, the posterior mode coincides with the MLE. *(ii)* in large samples, for any given prior, the posterior mode is very close to the MLE.

Therefore, if we face a situation with either no prior information and any sample size or any informative prior and a large sample, we can avoid Bayesian calculations and controversies by using a maximum likelihood estimator. This makes BAMLE easy to interpret, easy to apply and more flexible than BACE.

3.1 The Likelihood Function

The panel data methods employed in the aforementioned approaches only permit use of the within variation in the data, and therefore cannot exploit the information contained in regressors without time variation. This situation implies that we are not considering all the potential determinants of economic growth. For instance, some theories argue that geographic factors without time variation matter for growth. Moreover, as it is well-known, since assumption (A1) does not hold in dynamic panels, the within estimator of α is biased when T is small, as will be our case. Given the importance of this parameter -the convergence parameter- in the growth context, it is desirable to get an unbiased estimator of α. Given the Bayesian spirit of the approach, we propose here to use a maximum likelihood estimator - for a given model - which permits addressing the two drawbacks just described.

For a given model M_j we can write:

$$y_{it} = \alpha y_{it-\tau} + x_{it}'^{j}\beta^j + z_i^j\gamma^j + \eta_i + \zeta_t + v_{it}$$

Moreover, we can go further and assume[5]:

$$v_{it}|y_{it-1}...y_{i0}, x_i^j, z_i^j, \eta_i \sim N\left(0, \sigma_v^2\right) \tag{A3}$$

$$\eta_i|y_{i0}, x_i^j, z_i^j \sim N\left(\varphi y_{i0} + \delta^j \overline{x}_i^j, \sigma_\eta^2\right) \tag{A4}$$

Under assumptions (A3) and (A4) we can write the likelihood as[6]:

$$\begin{aligned} \log f\left(y_i|y_{i0}, x_i^j, z_i^j\right) \;\propto\; & -\frac{T-1}{2}\log\sigma_v^2 - \\ & -\frac{1}{2\sigma_v^2}\left(y_i^* - \alpha y_{i(-1)}^* - x_i^{*j}\beta^j\right)'\left(y_i^* - \alpha y_{i(-1)}^* - x_i^{*j}\beta^j\right) - \\ & -\frac{1}{2}\log\omega^2 - \frac{1}{2\omega^2}\left(\overline{y}_i - \alpha\overline{y}_{i(-1)} - \gamma^j z_i^j - \phi^j\overline{x}_i^j - \varphi y_{i0}\right)^2, \end{aligned} \tag{22}$$

where $\phi^j = \beta^j + \delta^j$, φ and ω^2 are the linear projection coefficients of $\overline{u}_i$ on $\overline{x}_i^j$ and y_{i0}, and y_i^*, $y_{i(-1)}^*$ and x_i^{*j} denote orthogonal deviations of y_i, $y_{i(-1)}$ and x_i^j respectively.

Thus, the Gaussian log-likelihood given y_{i0}, x_i^j and z_i^j can be decomposed into a within-group and a between-group component. This allows us to obtain an unbiased and consistent estimator for α (Alvarez and Arellano (2003)). Furthermore, the between-group component together with the orthogonality assumption between z_i^j and η_i allow for identification of γ^j.

We should emphasize that assumption (A4) implies that the regressors with and without temporal variation are treated differently. In the spirit of Hausman and Taylor (1981) but in a simpler framework, it is important to note that while the x's can be correlated with the unobservable fixed effect, the z's are independent. One interpretation is that, in addition to the traditional unobserved heterogeneity between countries given by the η_i term, there also exists a second type of fixed but observable heterogeneity given by the z_i variables. Moreover, both types of heterogeneity must be mutually uncorrelated. For instance, we may think about observable geographic factors such as land area, which are independent from unobservables of each country as could be the ability of its population. With the BAMLE approach, we will be able to conclude which observable fixed

[5] Note that all data will be cross-sectional de-meaned given the inclusion of time dummies.

[6] See Alvarez and Arellano (2003) for the demonstration in the pure autorregresive model. We add here additional exogenous explanatory variables with and without temporal variation.

factors are more important in promoting economic growth. This conclusion could also be obtained by using standard random effects estimation, but it is important to remark that with our approach we do not need to assume independence between the country specific effect and time varying regressors, which seems to be implausible in this context.

3.2 The BIC Approximation

Once we have specified the likelihood function of the data, we need a few more ingredients for the implementation of the BAMLE methodology. An essential one is the derivation of the integrated likelihood for a given model presented in equation (4). Various analytic and numerical approximations have been proposed to address this problem. In particular, we will make use of the Bayesian Information Criterion (BIC) approximation, which is both simple and accurate. The Schwarz criterion gives a rough approximation to the logarithm of the Bayes factor, which is easy to use and does not require evaluation of subjective prior distributions.

We can approximate the Bayes factor between models M_i and M_j, $B_{ij} = \frac{f(y|M_i)}{f(y|Mj)}$ such that (Raftery (1995)):

$$S = \log f\left(y|\widehat{\theta}_i, M_i\right) - \log f\left(y|\widehat{\theta}_j, M_j\right) - \frac{(k_i - k_j)}{2} \log(NT), \tag{23}$$

where $\widehat{\theta}_i$ is the MLE under M_i, k_i is the dimension of $\widehat{\theta}_i$, and NT is the sample size. As $NT \to \infty$, this quantity, often called the Schwarz criterion, satisfies:

$$\frac{S - \log B_{ij}}{\log B_{ij}} \to 0 \tag{24}$$

Minus twice the Schwarz criterion is often called the Bayesian information criterion (BIC):

$$BIC = -2S \approx -2 \log B_{ij}. \tag{25}$$

The relative error of $exp(S)$ in approximating B_{ij} is generally $O(1)$. Thus even for very large samples, it does not produce the correct value. On the other hand, we must keep in mind that in our approach, testing two competing hypothesis is not the final objective, and therefore we do not need the exact value of the Bayes factor. Instead we only need a rough interpretation of B_{ij} in a logarithmic scale such that[7]:

$2\log B_{ij}$	B_{ij}	Interpretation by the MC^3 algorithm
> 0	> 1	Strong evidence against M_j
< 0	< 1	Not strong evidence against any model

[7] This is the interpretation we need for the implementation of our approach with the MC^3 algorithm. See Computational Appendix for more details on the MC^3 algorithm.

Equation (24) shows that in large samples the Schwarz criterion is equivalent to the logarithm of the Bayes factor and therefore it should provide a reasonable indication of this evidence.

The value of BIC for model M_j denoted BIC_j, is the approximation to $-2logB_{0j}$ given by (25), where B_{0j} is the Bayes factor for model M_j against M_0 (which could be the null model with no independent variables). Moreover, we can manipulate the previous equations in the following manner:

$$\begin{aligned} B_{ij} &= \frac{f(y|M_i)}{f(y|M_j)} = \frac{\frac{f(y|M_i)}{f(y|M_0)}}{\frac{f(y|M_j)}{f(y|M_0)}} = \frac{B_{i0}}{B_{j0}} = \frac{B_{0j}}{B_{0i}}. \\ 2\log B_{ij} &= 2\left[\log B_{0j} - \log B_{0i}\right] = -BIC_j + BIC_i. \end{aligned}$$

In addition, we can rewrite equation (3) as:

$$\begin{aligned} P(M_j|y) &= \frac{f(y|M_j)P(M_j)}{\sum_{i=1}^{2^K} f(y|M_i)P(M_i)} = \\ &= \frac{\frac{f(y|M_j)}{f(y|M_h)} f(y|M_h)P(M_j)}{\sum_{i=1}^{2^K} \frac{f(y|M_i)}{f(y|M_h)} f(y|M_h)P(M_i)} = \\ &= \frac{B_{jh} f(y|M_h)P(M_j)}{\sum_{i=1}^{2^K} B_{ih} f(y|M_h)P(M_i)} = \\ &= \frac{\frac{1}{B_{0j}}P(M_j)}{\sum_{i=1}^{2^K} \frac{1}{B_{0i}}P(M_i)} = \frac{B_{j0}P(M_j)}{\sum_{i=1}^{2^K} B_{i0}P(M_i)}, \end{aligned} \tag{26}$$

where since $B_{00} = 1$, $BIC_0 = 0$, then $B_{j0} = \exp\left(\frac{1}{2}BIC_j\right)$.

Given the above, instead of integrating the marginal likelihood in (4), we will use the following result:

$$f(y|M_j) \propto \exp\left(\frac{1}{2}BIC_j\right), \tag{27}$$

and therefore:

$$P(M_j|y) = \frac{P(M_j)\exp\left(\frac{1}{2}BIC_j\right)}{\sum_{i=1}^{2^K} P(M_i)\exp\left(\frac{1}{2}BIC_i\right)}. \tag{28}$$

Furthermore, the posterior odds ($posterior\ odds = prior\ odds$ x $Bayes\ Factor$) becomes:

$$\frac{P(M_i|y)}{P(M_j|y)} = \frac{P(M_i)}{P(M_j)} \frac{\exp\left(\frac{1}{2}BIC_i\right)}{\exp\left(\frac{1}{2}BIC_j\right)}. \tag{29}$$

3.3 The Choice of Priors

Bayesian inference may be controversial because it requires specification of prior distributions which are subjectively chosen by the researcher. Moreover, Bayesian calculations may be extremely hard and computationally demanding when estimating millions of non-regular models[8].

Given the use of a maximum likelihood estimator and the BIC approximation, BAMLE avoids the need to specify a particular prior for the parameters of a given model.

As a result, for the implementation of BAMLE, the researcher only needs to specify priors on the model space. In particular, in an attempt to limit the effects of weakly held prior views, we suggest to employ the Binomial-Beta prior structure proposed by Ley and Steel (2007), as described in the previous section.

4 Data

A huge number of variables have been proposed as growth determinants in the cross-country literature, including variables with and without time variation. However, data for many of the former is not available over the entire sample period under consideration in this paper[9]. Since our main goal is to work with a panel data set, we limit our selection of time-varying variables to those for which data is available over the entire period 1960-2000.

In the construction of our data set, we have considered two different criteria. The first selection criterion derives from our aim of obtaining comparable results with the existing literature, and the second criterion comes from the fact that we need to work with a balanced panel.

With these restrictions, our data set includes a total of 35 variables (including the dependent variable, the growth rate of per capita GDP) for 73 countries and for the period 1960-2000. In order to avoid the problem of serial correlation in the transitory component of the disturbance term, we have split our sample in five year periods. Therefore we have eight observations for each country, that is to say, we have a sample of 584 observations.

Among the 19 regressors with temporal variation in our data set, there are both stock and flow variables. Following Caselli, Esquivel, and Lefort (1996), stock variables such as population and years of primary education are measured in the first year of each five-year period. On the other hand, flow variables such as population growth and investment rate are measured as five-year averages. Finally, as we focus on 5-year periods, $\tau = 5$ in all our estimated models.

4.1 Determinants of Economic Growth

The augmented Solow model can be taken as the baseline empirical growth model. It comprises four determinants of economic growth, initial income, rates of human

[8]I refer here to non-regular models as those for which closed-form solutions are not available when unsing informative priors.

[9]For instance, the fraction of GDP in mining and the fraction of Muslim population (both considered in Fernández *et. al.* (2001a) and Sala-i-Martin *et. al.* (2004)) are only available for the year 1960.

and physical capital accumulation, and population growth. We capture these growth determinants through the ratio of real investment to GDP, the stock of years of education and demographic variables such as life expectancy, the ratio of labor force to total population and population growth. In addition to those four determinants, Durlauf, Johnson, and Temple's (2005) 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. Due to data availability, our set of growth determinants is a subset of that identified by Durlauf, Johnson and Temple (2005). We consider the three broad variable categories below.

- Macroeconomic and external environment: A stable macroeconomic environment characterized by low and predictable inflation, sustainable budget deficits, and limited departure of the real exchange rate from its equilibrium level sends important signals to the private sector about the commitment and credibility of a country's authorities to efficiently manage their economy and increase the opportunity set of profitable investments. In this paper, the impact of macroeconomic stability is captured by the government consumption relative to GDP. Since the seminal work of Barro (1991), many authors have considered this ratio (g^c/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 lowers saving and growth through the distorting effects from taxation or government-expenditure programs. Moreover, following Easterly (1993) among others, we also consider the investment price level (*i.e.*, the PPP investment deflator) as a proxy for the level of distortions that exists in the economy. Finally, the trade regime/external environment is captured by the degree of trade openness, measured by imports plus exports as a share of GDP. Many authors such as Levine and Renelt (1992) and Frankel and Romer (1999) have considered this ratio. However, since this measure is sometimes criticized because it only captures the volume of trade and not the degree of openness as a proxy for distortions in trade policies, we also consider an alternative indicator, the SW openness index constructed by Sachs and Warner (1995). The objective is to conclude which measure of openness is a better (in the sense of more robust) proxy.

- Institutions and governance: The role of democracy and institutions in the process of economic growth has been the source of considerable research effort. In this paper we examine the hypothesis that political freedom and institutional quality are significant determinants of economic growth using political rights and civil liberties indices to measure the quality of institutions and capture the occurrence of free and fair elections and decentralized political power. Kormendi and Meguire (1985), Barro (1991), Barro and Lee (1994) and Sala-i-Martin (1997) among others considered these two indices as proxies of the quality of institutions and governance.

- Geography and fixed factors: Following Sachs and Warner (1997) and Bloom and Warner (1998), there is an influential view arguing that differences in natural endowments, such as climatic conditions can account for income differences accross countries. Very closely related, another view stresses market access (remoteness) in explaining spatial variation in economic activity, as emphasized in the literature on new economic geography following Krugman (1991). In order to examine the extent to which geography matters for growth, we use a variety of geographic indicators such as the percentage of land area in the geographical tropics or the fraction of population in geographical tropics. On the other hand, as proxies for remoteness we use, among others, the minimal distance to New York, Rotterdam or Tokio, the fraction of land area near navigable water and a dummy for landlocked countries. Finally, other fixed but not geographic factors such as the active participation in conflicts during the sample period[10] (war dummy) or the timing of independence, may have an effect on economic growth as pointed out by Barro and Lee (1994) and Gallup et. al. (2001) respectively.

A list of variables with their corresponding description and sources can be found in the Data Appendix, as well as the list of countries included in the sample.

5 Results

Table 1 reports the posterior inclusion probability of the 19 regressors with time variation included in our data set after applying BACE-SDM and BMA-FLS procedures. The table highlights the sensivity of the results to the different prior assumptions. Concretely, comparison of columns 1 and 3, and 2 and 4, shows that with fixed ξ different assumptions about the prior mean model size, $m = 5$ or $m = K/2$, generate quite different posterior inclusion probabilities. More specifically, when we do not penalize larger models in any way – that is to say, when we employ $m = K/2$ instead of $m = 5$ in the BACE-SDM approach (columns 3 and 1 respectively) – the posterior inclusion probabilities are higher. On the other hand, when we do penalize bigger models in both ways employing $m = 5$ in the BMA-FLS approach (column 2), the posterior inclusion probabilities are smaller. This also highlights the "fortuitous robustness" which emerges when we compare the BMA-FLS and BACE-SDM's results in columns 1 and 4, that is to say, different prior assumptions on model size have substantial effects on the results. Furthermore, analyzing columns 5 to 8 of Table 1, we can conclude that the effects of prior assumptions on model size are much less important in the case of random ξ. Moreover, the last row of the table indicates that expected model size should be close to 5 in the panel data framework.

[10] Given data availability and the requirement of a balanced panel, we follow Barro and Lee (1994) and use a dummy variable for countries that participated in at least one external war over the period 1960-1990. Then, this variable is considered here as fixed over the sample period.

Table 2 shows the posterior inclusion probability, the posterior mean and the posterior standard error for the parameters corresponding to the 19 variables of our data set with time variation when we apply the BACE-SDM and BMA-FLS approaches in a panel data context. These results are based on the whole sample, that is, 73 countries for the period 1960-2000. The main conclusion from the table is that, in addition to initial GDP, there are several covariates which appear robustly related to economic growth. However, we defer our main conclusions to Table 4 below.

In Tables 3 and 5, we follow the methodology employed by Ciccone and Jarocinski (2005). Employing the same sample period for both sources of income data[11], we can assess the sensitivity to changes in data source of the results in terms of posterior inclusion probability and posterior mean. The measures shown are self-explanatory: the results are considerably less sensitive to differences in income data source than found in the previous literature, at least for the comparison between World Bank and Penn World Table income data[12]. In order to further explore this issue, we redo the sensivity analysis using the BACE-SDM approach without considering the panel structure of the data. By doing this, the only difference vis-a-vis Ciccone and Jarocinsky (2005) is the number of regressors considered in the exercise. While they consider 67 potential explanatory variables, we consider 34. Looking at the results, presented in Table 6, we can see that the sensivity with $K = 34$ is much smaller than with $K = 67$. Therefore, we conclude that the number of potential explanatory variables under consideration is critical for the sensivity of the results to changes in the source of international income data used. Concretely, the fewer the regressors, the smaller the sensivity.

Results when applying the BAMLE Approach with PWT 6.2 income data for the whole period are summarized in Table 4. Addionally to initial GDP, a fair number of regressors could be considered as robust determinants of economic growth accordingly to the Bayesian robustness check used in the approach. The most conclusive evidence is for investment price, distance to major world cities and political rights. All three regressors affect growth with the expected sign: in particular, as found by Easterly (1993), a low level of distorsions in the economy (*i.e.* lower investment price) would promote economic growth. A better geographic situation, (*i.e.* a better access to international markets) is also an important growth enhancing factor as argued by Krugman (1991). Finally, in contrast to Barro (1991) but in line with Sala-i-Martin (1997), a higher level of democracy (*i.e.* a lower value of the variable measuring restrictions on political rights) is found to be positively related to higher growth rates. On the other hand, since their posterior inclusion probability is higher than their prior inclusion probability, many other variables such as demographic indicators, a measure of trade openness, the dummy for landlocked countries, the investment share, the civil liberties index and the government share can be considered as robust determinants of economic growth. Finally, there is one regressor, life expectancy, that poses a puzzle. In

[11] Note that WDI 2005 income data only covers the period 1975-2000.

[12] See Ciccone and Jarocinski (2005) for more details on the cross-country context.

spite of having the highest posterior inclusion probability, we think it cannot be viewed as robust because its posterior standard error is bigger than its posterior mean. Despite the inclusion of country-specific effects correlated with the time-varying variables, it is important to note that all these results must be interpreted with some caution, since they assume that the x's variables are strictly exogenous with respect to the transitory component of the disturbance, which might not be a valid assumption in this context.

It is worth mentioning that the posterior mean conditional on inclusion of the lagged dependent variable (initial GDP) in Table 4 implies a rate of conditional convergence of $\lambda = 0.006$. This suggest that after controlling for model uncertainty and other potential inconsistencies afecting the lagged dependent variable (arising from omitted variable and endogeneity biases), the estimated rate of convergence is surprisingly similar to the standard cross-section finding[13].

6 Concluding Remarks

In spite of a huge amount of empirical research, the drivers of economic growth are not well understood. This paper attempts to provide insights on the growth puzzle by searching for robust determinants of economic growth. We propose a Bayesian Averaging of Maximum Likelihood Estimates (BAMLE) method in a panel data framework to determine which variables are significantly related to growth. Similarly to the BACE approach, our method is more appealing than a standard Bayesian Model Averaging since it does not require the specification of prior distributions for the parameters of every model under consideration, and it involves only one hyper-parameter, expected model size m. Moreover, the BAMLE approach is more flexible than BACE and it introduces two improvements with respect to previous model-averaging and robustness-checking methods applied to empirical growth regressions: (i) it adressess the problem of inconsistent empirical estimates by using a dynamic panel estimator, and (ii) it minimizes the impact of prior assumptions about the only hyper-parameter in the approach. An additional methodological conclusion of the paper is that the list of growth determinants emerging from a set of 34 potential explanatory variables is less sensitive to the use of alternative sources of international income data than in the case of other papers which considered a larger number of potential regressors. Therefore, we conclude that the fewer the potential growth determinants considered, the smaller the sensivity to changes in growth data.

The empirical findings suggest that country specific effects correlated with other regressors play an important role since the list of robust growth determinants is not the same when we do not take into account their presence. Our results indicate that once model uncertainty and other potential inconsistencies are accounted for, there exist economic, institutional, geographic and demographic factors that robustly affect growth. The most robust determinants are investment

[13]See for example Mankiw, Romer and Weil (1992, Table 4).

price, distance to major world cities and political rights. Other variables which can be considered as robust include demographic factors (population growth, urban population and population), geographical dummies (such as the dummy for landlocked countries), measures of openness and civil liberties, and macroeconomic indicators such as the investment share of GDP and the ratio of government consumption to GDP. On the other hand, our empirical estimate of the rate of convergence, after controlling for both model uncertainty and endogeneity, is surprisingly similar to that commonly found in cross-section studies.

As a final remark, it is worth mentioning that the dynamic panel estimator proposed in this paper addresses the endogeneity of regressors with time variation with respect to the permanent component of the error term as well as the endogeneity of the lagged dependent variable with respect to the transitory component of the error term. However, many other regressors such as the labor force or the investment share should ideally be considered as predetermined instead of strictly exogenous with respect to the transitory component of the error term, and this point remains unresolved in the BMA context. Hence, the estimates might change under less stringent exogeneity assumptions. This issue is left for future research.

A Appendix

A.1 Computational Appendix

For the implementation of the empirical approaches described in the paper, we need to resort to the algorithms proposed in the literature because of the extremely large number of calculations required for obtaining the posterior mean and variance described in equations (6) and (7). This is because the number of potential regressors determines the number of models under consideration, for example, in our case, with $K = 35$ potential regressors, the number of models under consideration is 3.4x10^{10}. These algorithms carry out Bayesian Model Averaging without evaluating every possible model.

Concretely, for the BACE, BMA and BAMLE approaches we have made use of the Markov Chain Monte Carlo Model Composition (MC^3) algorithm proposed by Madigan and York (1995), which generates a stochastic process that moves through model space. The idea is to construct a Markov chain of models $\{M(t), t = 1, 2, ...\}$ with state space Ξ. If we simulate this Markov chain for $t = 1, ..., N$, then under certain regularity conditions, for any function $h(M_i)$ defined on Ξ, the average

$$\widehat{H} = \frac{1}{N}\sum_{t=1}^{N} h\left(M\left(t\right)\right)$$

converges with probability 1 to $E\left(h\left(M\right)\right)$ as $N \rightarrow \infty$. To compute (6) in this fashion, we set $h(M_i) = E(\theta|M_i, y)$.

To construct the Markov chain, we define a neighborhood $nbd(M)$ for each $M \in \Xi$ that consists of the model M itself and the set of models with either one variable more or one variable fewer than M. Then, a transition matrix $\mathbf{q}$ is defined by setting $\mathbf{q}(M \rightarrow M') = 0 \ \forall \ M' \notin ndb(M)$ and $\mathbf{q}(M \rightarrow M')$ constant for all $M' \in ndb(M)$. If the chain is currently in state M, then we proceed by drawing M' from $\mathbf{q}(M \rightarrow M')$. It is the accepted with probability

$$\min\left\{1, \frac{\Pr\left(M'|y\right)}{\Pr\left(M|y\right)}\right\}$$

Otherwise, the chain stays in state M[14].

After some experimentation with generated data, we verify the proper convergence properties of our Gauss code which implements the described MC^3 algorithm.

[14]Koop (2003) is a good reference for the reader interested in developing a deeper understanding of the MC^3 algorithm.

A.2 Data Appendix

Table A1: Variable Definitions and Sources

Variable	Source	Definition
Dependent Variable	PWT 6.2	Growth of GDP per capita over 5-year periods (2000 US dollars at PPP)
Initial GDP	PWT 6.2	Logarithm of initial real GDP per capita (2000 US dollars at PPP)
Population Growth	PWT 6.2	Average growth rate of population
Population	PWT 6.2	Population in thousands of people
Trade Openness	PWT 6.2	Export plus imports as a share of GDP
Government Share	PWT 6.2	Government consumption as a share of GDP
Investment Price	PWT 6.2	Average investment price level
Labor Force	PWT 6.2	Ratio of workers to population
Consumption Share	PWT 6.2	Consumption as a share of GDP
Investment Share	PWT 6.2	Investment as a share of GDP
Urban Population	WDI 2005	Fraction of population living in urban areas
Population Density	WDI 2005	Population divided by land area
Life Expectancy	WDI 2005	Life expectancy at birth
Population under 15	Barro and Lee	Fraction of population younger than 15 years
Population over 65	Barro and Lee	Fraction of population older than 65 years
Primary Education	Barro and Lee	Stock of years of primary education
Secondary Education	Barro and Lee	Stock of years of secondary education
Political Rights	Freedom House	Index of political rights from 1 (highest) to 7
Civil Liberties	Freedom House	Index of civil liberties from 1 (highest) to 7
Malaria	Gallup et. al.	Fraction of population in areas with malaria
Navigable Water	Gallup et. al.	Fraction of land area near navigable water
Landlocked Country	Gallup et. al.	Dummy for landlocked countries
Air Distance	Gallup et. al.	Logarithm of minimal distance in km from New York, Rotterdam, or Tokio
Tropical Area	Gallup et. al.	Fraction of land area in geographical tropics

Notes:
1. PWT 6.2 refers to Penn World Table 6.2
2. WDI 2005 refers to World Development Indicators 2005 from The World Bank

Table A1 - Continued

Variable	Source	Definition
Tropical Pop.	Gallup et. al.	Fraction of population in geographical tropics
Land Area	Gallup et. al.	Area in km^2
Independence	Gallup et. al.	Timing of national independence measure: 0 if before 1914; 1 if between 1914 and 1945; 2 if between 1946 and 1989 and 3 if after 1989
Socialist	Gallup et. al.	Dummy for countries under socialist rule for considerable time during 1950 to 1995
Climate	Gallup et. al.	Fraction of land area with tropical climate
War Dummy	Barro and Lee	Dummy for countries that participated in external war between 1960 and 1990
SW Openness Index	Sachs, Warner	Index of trade opennes from 1 (highest) to 0
Europe		Dummy for EU countries
Sub-Saharan Africa		Dummy for Sub-Saharan African countries
Latin America		Dummy for Latin American countries
East Asia		Dummy for East Asian countries

Table A2: 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
Dominican 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

[1] Alvarez, J. and M. Arellano (2003), "The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators" *Econometrica*, Vol. 71, No. 4. pp. 1121-1159.

[2] Barro, R. (1991), "Economic Growth in a Cross Section of Countries" *Quarterly Journal of Economics*, 106, 2, 407-43

[3] Barro, R. and J.-W. Lee (1994), "Sources of Economic Growth" *Carnegie-Rochester Conference Series on Public Policy*, 40, 1-57.

[4] Bloom, D. and J. Sachs. (1998), "Geography, Demography and Economic Growth in Africa." *Brookings Papers on Economic Activity*, 207—73.

[5] Brock, W. and S. Durlauf (2001), "Growth Empirics and Reality" *World Bank Economic Review*, 15, 2, pp. 229-272.

[6] 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, pp. 363-389.

[7] Ciccone, A. and M. Jarocinski (2005), "Determinants of Economic Growth: Will Data Tell?" Unpublished manuscript.

[8] Durlauf, S., P. Johnson and J. Temple (2005), "Growth Econometrics" In P. Aghion and S.N. Durlauf, eds., *Handbook of Economic Growth*, Volume 1A, pp. 555-677, Amsterdam, North-Holland.

[9] Easterly, W. (1993), "How Much Do Distortions Affect Growth?" *Journal of Monetary Economics*, 32, 2, 187-212.

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

[11] Fernandez, C., E. Ley and M. Steel (2001b), "Benchmark Priors for Bayesian Model Averaging" *Journal of Econometrics*, 100, pp. 381-427.

[12] Frankel, J. and D. Romer (1999), "Does Trade Cause Growth?," *American Economic Review*, 89, 3, 379-399.

[13] Gallup, J., A. Mellinger and J. Sachs (2001) "Geography Datasets" *Center for International Development at Harvard University (CID)*

[14] Hausman, J. and W. Taylor, (1981) "Panel Data and Unobservable Individual Effects" *Econometrica*, Vol. 49, No. 6. pp. 1377-1398.

[15] Islam, N. (1995), "Growth Empirics: A Panel Data Approach" *The Quarterly Journal of Economics*, Vol. 110(4), pp. 1127-70

[16] Kass, R. and A. Raftery (1995), "Bayes Factors" *Journal of the American Statistical Association*, Vol. 90, No 430, pp. 773-795.

[17] Kormendi, R. and P. Meguire (1985), "Macroeconomic Determinants of Growth: Cross Country Evidence," *Journal of Monetary Economics*, 16, 2, 141-63.

[18] Krugman, P. (1991), "Increasing Returns and Economic Geography." *Journal of Political Economy*, 99(3): 483—99.

[19] León-González, R. and D. Montolio (2004), "Growth, Convergence and Public Investment: A BMA Approach" *Applied Economics*, 36, pp. 1925-36.

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

[21] Ley, E. and M. Steel (2007), "On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression" Unpublished manuscript.

[22] Madigan, D. and J. York (1995), "Bayesian Graphical Models for Discrete Data" *International Statistical Review*, 63, pp. 215-232.

[23] Mankiw, N., D. Romer and D. Weil (1992), "A Contribution to the Empirics of Economic Growth" *Quarterly Journal of Economics*, 107, pp. 407-437.

[24] Masanjala, W. and C. Papageorgiou (2005), "Rough and Lonely Road to Prosperity: A reexamination of the sources of growth in Africa using Bayesian Model Averaging" Unpublished manuscript.

[25] Raftery, A. (1995), "Bayesian Model Selection in Social Research" *Sociological Methodology*, Vol. 25, pp. 111-163.

[26] Raftery, A. (1996), "Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models" *Biometrika*, Vol. 83, No. 2.pp. 251-266.

[27] Sachs, J. and Warner, A. (1995), "Economic Reform and the Process of Economic Integration", *Brookings Papers of Economic Activity*, pp.1-95.

[28] Sachs, J. and Warner, A. (1997), "Natural Resource Abundance and Economic Growth", *CID at Harvard University.*

[29] Sala-i-Martin, X. (1997), "I Just Ran Two Million Regressions" *American Economic Review*, Vol. 87, No. 2. pp. 178-183. Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association.

[30] 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*, Vol. 94, No. 4. pp. 813-835.

[31] Schwarz, G. (1978), "Estimating the Dimension of a Model" *The Annals of Statistics*, Vol. 6, No. 2. pp. 461-464.

[32] Tsangarides, C. (2004), "A Bayesian Approach to Model Uncertainty" *IMF Working Paper* WP/04/68.

[33] Tsangarides, C. (2005), "Growth Empirics Under Model Uncertainty: Is Africa Different?" *IMF Working Paper* WP/05/18.

Tables

Table 1: Posterior Inclusion Probability of the Regressors

Variable	ξ Fixed				ξ Random			
	$m = 5$		$m = K/2$		$m = 5$		$m = K/2$	
	SDM	FLS	SDM	FLS	SDM	FLS	SDM	FLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Initial GDP	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Population	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Population under 15	0.950	0.961	0.937	0.953	0.953	0.965	0.949	0.963
Investment Share	0.826	0.847	0.783	0.835	0.822	0.841	0.816	0.843
Urban Population	0.651	0.392	0.781	0.596	0.608	0.358	0.638	0.387
Consumption Share	0.305	0.100	0.682	0.229	0.303	0.088	0.351	0.099
Trade Opennes	0.287	0.106	0.656	0.218	0.289	0.094	0.336	0.103
Government Share	0.237	0.064	0.549	0.173	0.231	0.058	0.273	0.068
Investment Price	0.222	0.088	0.376	0.176	0.206	0.083	0.229	0.092
Population Density	0.031	0.013	0.061	0.024	0.029	0.011	0.033	0.013
Labor Force	0.029	0.013	0.064	0.022	0.028	0.010	0.033	0.012
Primary Education	0.026	0.010	0.061	0.023	0.026	0.009	0.030	0.010
Civil Liberties	0.023	0.007	0.053	0.017	0.022	0.006	0.025	0.008
Population Growth	0.018	0.005	0.050	0.013	0.019	0.005	0.022	0.005
Life Expectancy	0.018	0.006	0.051	0.013	0.019	0.005	0.023	0.006
Malaria	0.020	0.005	0.043	0.014	0.018	0.006	0.021	0.006
Population over 65	0.017	0.005	0.044	0.013	0.018	0.004	0.021	0.006
Secondary Education	0.017	0.005	0.046	0.012	0.017	0.005	0.020	0.005
Political Rights	0.016	0.005	0.044	0.012	0.016	0.004	0.020	0.005
Prior Mean Model Size	5	5	9	9	5	5	9	9
Post. Mean Model Size	5.69	4.63	7.28	5.34	5.62	4.55	5.83	4.63

Column heading SDM refers to the BACE-SDM Approach in a panel data context.
Column heading FLS refers to BMA-FLS approach in a panel data context.

Table 2: SDM-FLS Approaches in a Panel Data Context
with PWT 6.2 Income Data 1960-2000*

Variable	Posterior Inclusion Probability		Posterior Mean		Posterior Standard Error	
	SDM	FLS	SDM	FLS	SDM	FLS
Initial GDP	1.000	1.000	-0.271	-0.265	0.029	0.030
Population	1.000	1.000	0.918	0.905	0.176	0.176
Population under 15	0.953	0.965	-1.122	-1.183	0.287	0.279
Investment Share	0.822	0.841	0.343	0.351	0.097	0.095
Urban Population	0.608	0.358	-0.426	-0.433	0.147	0.147
Consumption Share	0.303	0.088	-0.210	-0.202	0.068	0.091
Trade Opennes	0.289	0.094	0.102	0.100	0.028	0.046
Government Share	0.231	0.058	-0.336	-0.315	0.140	0.149
Investment Price	0.206	0.083	-0.031	-0.033	0.014	0.014
Population Density	0.029	0.011	0.042	0.063	0.054	0.057
Labor Force	0.028	0.010	0.225	0.363	0.415	0.477
Primary Education	0.026	0.009	-0.169	-0.194	0.179	0.186
Civil Liberties	0.022	0.006	-0.044	-0.047	0.060	0.060
Population Growth	0.019	0.005	-0.488	-0.317	1.156	1.091
Life Expectancy	0.019	0.005	0.063	-0.011	0.241	0.250
Malaria	0.018	0.006	0.010	0.013	0.024	0.026
Population over 65	0.018	0.004	-0.220	-0.200	0.824	0.801
Secondary Education	0.017	0.005	-0.051	-0.034	0.186	0.191
Political Rights	0.016	0.004	-0.009	-0.004	0.048	0.049

*All results presented in this Table are based on prior assumptions $m = 5$ and ξ Random. The results with $m = K/2$ are not presented here for the sake of brevity, but they were practically identical.

Table 3: Sensivity Analysis PWT 6.2 vs. WDI 2005
with SDM-FLS Approaches in a Panel Data Context

Variable	MAX/MIN Posterior Inclusion Probability		[MAX-MIN]/ABS(MIN) Posterior Mean	
	SDM	FLS	SDM	FLS
Initial GDP	1.000	1.000	0.113	0.114
Population	1.002	1.007	0.057	0.063
Population under 15	1.024	1.021	0.128	0.133
Investment Share	1.482	1.647	0.328	0.307
Urban Population	2.758	3.677	0.204	0.204
Consumption Share	1.064	1.139	0.110	0.129
Trade Opennes	1.933	2.580	0.352	0.341
Government Share	1.104	1.180	0.031	0.038
Investment Price	1.644	1.610	0.049	0.034
Population Density	1.068	1.222	2.095	2.445
Labor Force	1.182	1.184	2.022	1.931
Primary Education	2.420	2.700	0.959	0.986
Civil Liberties	4.308	4.536	0.537	0.521
Population Growth	4.062	6.144	4.286	6.333
Life Expectancy	1.265	1.335	6.041	3.465
Malaria	1.004	1.129	0.728	0.833
Population over 65	5.651	6.582	0.880	0.856
Secondary Education	1.171	1.070	0.058	0.020
Political Rights	1.557	1.650	0.409	0.410
Average	1.932	2.232	1.020	1.009
Median	1.265	1.335	0.352	0.341

Table 4: BAMLE Approach with PWT 6.2 Income Data 1960-2000*

Variable	Posterior Inclusion Probability	Posterior Mean	Posterior Standard Error
Initial GDP	1.000	-0.033	0.035
Life Expectancy	1.000	0.145	0.287
Investment Price	0.863	-0.049	0.015
Air Distance	0.759	-0.962	0.381
Political Rights	0.722	-0.053	0.013
Population Growth	0.688	-1.082	1.081
Urban Population	0.650	-0.475	0.163
Population	0.639	0.602	0.201
Trade Openness	0.467	0.056	0.020
Landlocked Country	0.320	-0.346	0.359
Investment Share	0.238	0.271	0.105
Civil Liberties	0.176	0.048	0.017
Government Share	0.161	-0.160	0.148
Latin America	0.147	0.038	0.015
Population Density	0.087	-0.014	0.081
East Asia	0.073	-0.012	0.006
Consumption Share	0.057	0.036	0.062
Navigable Water	0.057	0.043	0.026
Europe	0.052	-0.036	0.018
Tropical Area	0.034	-0.252	0.201
Sub-Saharan Africa	0.029	0.027	0.021
Climate	0.028	-0.014	0.013
Primary Education	0.028	0.024	0.022
Tropical Pop.	0.025	-0.144	0.212
Labor Force	0.023	0.028	0.394
Population over 65	0.022	-0.012	0.018
SW Openness Index	0.018	-0.033	0.069
Land Area	0.017	0.021	0.056
War Dummy	0.017	0.001	0.019
Population under 15	0.017	0.010	0.012
Secondary Education	0.017	-0.008	0.016
Independence	0.016	-0.002	0.015
Socialist	0.016	-0.009	0.013
Malaria	0.013	0.001	0.012

*All results presented in this Table are based on prior assumptions $m = 5$ and ξ Random.

Table 5: Sensivity Analysis PWT 6.2 vs. WDI 2005 with BAMLE Approach

Variable	MAX/MIN Posterior Inclusion Probability	[MAX-MIN]/ABS(MIN) Posterior Mean
Initial GDP	1.000	0.140
Life Expectancy	1.000	2.038
Investment Price	1.182	0.063
Air Distance	4.019	0.789
Political Rights	1.149	0.061
Population Growth	2.102	2.526
Urban Population	3.645	0.486
Population	1.215	0.221
Trade Openness	3.268	3.667
Landlocked Country	1.011	0.010
Investment Share	4.289	0.368
Civil Liberties	2.721	0.182
Government Share	2.895	1.964
Latin America	1.412	0.161
Population Density	1.738	0.878
East Asia	1.048	1.286
Consumption Share	1.792	0.456
Navigable Water	6.784	0.088
Europe	2.308	0.667
Tropical Area	1.154	0.878
Sub-Saharan Africa	1.042	0.343
Climate	2.079	0.346
Primary Education	1.059	5.000
Tropical Pop.	1.286	2.935
Labor Force	1.571	0.448
Population over 65	6.000	0.870
SW Openness Index	4.333	0.424
Land Area	1.882	0.518
War Dummy	1.000	0.222
Population under 15	1.643	0.250
Secondary Education	1.067	1.000
Independence	1.308	0.833
Socialist	1.571	1.304
Malaria	1.125	0.556
Average	2.138	0.940
Median	1.571	0.502

Table 6: Sensivity Analysis PWT 6.2 vs. WDI 2005
with BACE Approach and K=34*

	MAX/MIN Posterior Inclusion Probability		[MAX-MIN]/ABS(MIN) Posterior Mean	
	POOLED	OLS	POOLED	OLS
Average	2.889	1.892	1.513	1.306
Median	1.288	1.402	0.422	0.329

*Results based on BACE-SDM approach with cross-section OLS and panel POOLED estimation for a given model without considering the presence of unobservable fixed effects.