WORKING PAPERS

Jesus Crespo Cuaresma
Gernot Doppelhofer

Nonlinearities in Cross-Country Growth Regressions: A Bayesian Averaging of Thresholds (BAT) Approach

September 2006

Working Paper No: 0608

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

Nonlinearities in Cross-Country Growth Regressions: A Bayesian Averaging of Thresholds (BAT) Approach*

Jesus Crespo Cuaresma†
University of Vienna

Gernot Doppelhofer‡
University of Cambridge

September 2006

Abstract

We propose a framework for assessing the existence and quantifying the effect of threshold effects in cross-country growth regressions in the presence of model uncertainty. The method is based on Bayesian model averaging techniques and generalizes the Bayesian Averaging of Classical Estimates (BACE) method put forward by Sala-i-Martin, Doppelhofer, and Miller (2004). We apply the method presented in this paper to a set of 21 variables that have been found to be robustly related to economic growth in a cross-section of 88 countries. We find no evidence of robust threshold effects generated by the initial level of GDP per capita. However, we find that the proportion of years a country has been open to trade is an important source of nonlinear effects on economic growth.

Keywords: Model Uncertainty, Threshold Estimation, Non-Linearities, Determinants of Economic Growth

JEL Classifications: C11, C15, O20, O50

---

*We thank Theo Eicher, Chris Papageorgiou and Hashem Pesaran for helpful discussions. All errors are our own.

†Department of Economics, University of Vienna, Bruennerstrasse 72, 1210 Vienna, Austria. Email: jesus.crespo-cuaresma@univie.ac.at, Tel: +43 +1 4277 37477, Fax: +43 +1 4277 37490.

‡Faculty of Economics and Trinity College, University of Cambridge, Cambridge CB3 9DD, UK. Email: gd237@cam.ac.uk, Tel: +44 1223 335264, Fax: +44 1223 335475.

1 Introduction

Following the influential contributions of Kormendi and Meguire (1985) and Barro (1991), the empirical growth literature has used cross-country regressions to identify variables that are robustly (partially) correlated to growth of per capita GDP. Many different economic, social and political variables have been proposed as important determinants of economic growth. Durlauf and Quah (1999), for instance, name more than eighty variables that have been included at least once in a cross-country growth regression. Brock and Durlauf (2001) refer to this problem as the "open-endedness" of theories of economic growth.

Levine and Renelt (1992) gave a first assessment of the robustness of growth determinants by applying a version of Leamer (1983)'s extreme bounds analysis. Levine and Renelt checked for robustness when changing the set of conditioning variables and concluded that almost no variable used by Kormendi and Meguire (1985) and Barro (1991) is robustly correlated with average GDP per capita growth. Sala-i-Martin (1997a, 1997b), however, considers that the robustness test implied by extreme bounds analysis is too strong for any variable to pass it in the framework of empirical growth research, and proposed to analyze the entire distribution of estimates of the partial correlation between a given variable and long-run growth. Adopting such an approach, Sala-i-Martin (1997a, 1997b) attaches a "confidence level" (in terms of the probability mass on one side of zero in the empirical distribution of the estimate of the partial correlation) to each variable, and proposes to consider those variables with a confidence level of 95% or more as robustly correlated with long-run growth. Using this method, the conclusion is that there exists a considerable number of economic, political and demographic variables that are actually (partially) correlated to growth in a robust fashion.

The methods used to assess the robustness of covariates in growth regressions used by Levine and Renelt (1992) and Sala-i-Martin (1997a, 1997b) rely on models of a given size, so model uncertainty concerning the number of variables that should be included in the growth regression is not considered. Bayesian model averaging methods allow to account for model uncertainty both in the size of the model and in the choice of explanatory variables. Sala-i-Martin, Doppelhofer and Miller (2004) - henceforth SDM (2004) - introduce an alternative approach, Bayesian Averaging of Classical Estimates, BACE, that builds upon Bayesian model averaging without needing to specify prior distributions for all parameters in the econometric specification.[1] The method can be applied simply by repeated OLS estimations and presents a tractable setting aimed at accounting for model uncertainty in linear growth regressions. The results in SDM (2004) are in line with Sala-i- Martin (1997a, 1997b), indicating that there is a sizable group of variables which are robust explanatory factors for economic growth.

All the methods named above approach the issue of model uncertainty in growth regressions under the assumption that the relationship between the explanatory variables and the growth rate is *linear*. This essentially implies that the effect associated with a particular variable is constant across subsamples of the data used. Various deviations from the linear paradigm have been tested in the empirical literature and there is ample evidence of parameter heterogeneity, multiple regimes and threshold nonlinearities in cross country growth regressions (see e.g. Durlauf and Johnson, 1995, Durlauf, Kourtellos and Minkin, 2001, Masanjala and Papageorgiou, 2004, or Papageorgiou, 2002).[2] Many theoretical growth models deliver multiple steady states (e.g. Azariadis and Drazen, 1990). Masanjala and Papageorgiou (2004) have explicitly model nonlinearities in the aggregate production function. Finally, issues such as "poverty traps" and other threshold effects have been very influential in economic policy-making, motivating for example some of the "Millenium Goals" proposed by the United Nations.[3] The existence and economic importance of nonlinearities and threshold effects among determinants of economic growth plays thus a major role in the present policy discussion on global development strategies.

In this paper we explicitly allow for non-linearities in the form of level-dependent parameter heterogeneity as usually specified by threshold models (see Hansen, 1996, 2000). We allow for uncertainty over possible threshold effects and associated threshold observations by extending the BACE method of SDM (2004) and estimating the posterior distribution of these quantities of interest. We propose a method for estimating threshold values under model uncertainty based on the inspection of the posterior inclusion probability of the threshold parameter. Note that the distribution of threshold effects and interactions are calculated by averaging over many

[1]See also Fernández, Ley and Steel (2001) for an approach to robustness evaluation in cross-country growth regressions using Bayesian model averaging.

[2]Crespo Cuaresma (2002) presents a robustness exercise where threshold nonlinearity is explicitly accounted for but model size uncertainty is not dealt with.

[3]For an interesting debate on this issue see the interchange between Jeffrey Sachs and William Easterly at `http://www.nyu.edu/fas/institute/dri/Easterly/index.html`.

possible specifications and are therefore not conditional on a particular model. The resulting inference and policy analysis is therefore taking into account uncertainty over models, including nonlinear effects.

The paper contributes to the literature on the empirics of economic growth nonlinearities as follows: First, our proposed method allows the estimation of the entire posterior distribution of thresholds and associated nonlinear effects. We only need to specify a set of candidate threshold variables (motivated by the literature on growth nonlinearities discussed above) and prior parameters for the expected number of explanatory and threshold variables being present in the model. Second, we show that once we allow for uncertainty over the number of threshold variables and threshold observations, there is a relatively small set of robust nonlinear effects. In particular, conditioning on the *Number of Years an Economy Has Been Open* affects the size and significance of the effect of some other determinants of growth, whereas *Initial Income* appears to play a much less prominent role as a threshold variable when allowing for uncertainty over the number of variables causing nonlinearities. This result can be contrasted with the vast evidence of (model specific) nonlinearities found in earlier studies. Third, a technical contribution of the paper is to extend the BACE sampling method to the estimation of the distribution of nonlinear effects and associated thresholds. A key role is played by the specification of priors of inclusion of threshold variables and thresholds. We also extend the "stratified" sampler proposed by SDM (2004) to the case of threshold regressions.

The paper is organized as follows. Section two presents the methodology proposed to account for threshold nonlinearity in cross-country growth regressions in the presence of model uncertainty, which we call Bayesian Averaging of Thresholds (BAT). Section three reports the results of the robustness analysis for a dataset formed by the 21 variables that SDM (2004) find robust in their analysis and two potential threshold variables: the initial level of GDP per capita and the proportion of years that the economy has been open. Section four concludes and presents further paths of research.

2 Bayesian Averaging of Thresholds (BAT)

2.1 Thresholds and model uncertainty: BAT

The BAT approach is aimed at evaluating the existence and robustness of nonlinearities in regressions with model uncertainty. It is a generalization of the BACE approach in SDM (2004) which allows for threshold effects of certain variables on the regression parameters.

Consider a set of variables that are potentially related to growth, $\mathbf{X}$, and a set of variables that are potentially causing threshold-nonlinearity in the growth regression, $\mathbf{Z}$. $\mathbf{Z}$ may or may not be a subset of $\mathbf{X}$. The stylized nonlinear model we are considering is

$$\gamma = \alpha + \sum_{k=1}^{n} \beta_k x_k + \sum_{j=1}^{m} \left[(\alpha_j^* + \sum_{k=1}^{n} \beta_{jk}^* x_k) \mathsf{I}(z_j \leq \tau_j) \right] + \varepsilon, \tag{1}$$

where γ is a vector of T observations of growth rates of GDP per capita, $x_1, \ldots, x_n \in \mathbf{X}$, $z_1, \ldots, z_p \in \mathbf{Z}$, $\mathsf{I}(\cdot)$ is the indicator function, taking value one if the argument is true and zero otherwise and ε is an error term assumed uncorrelated across cross-sectional units and with constant variance σ^2. There are therefore m variables inducing nonlinearity in (1) and for simplicity we assume that the nonlinearity which is induced by variable z_i is independent from the regime in which an observation is according to another threshold variable z_j, for $i \neq j$. Although the BAT method can be generalized in a straightforward manner to the setting with dependent nonlinearities, this assumption avoids having to use cross-products of indicator functions in (1), which would increase the computational time of the procedure significantly.

Since we are explicitly dealing with model uncertainty, n and m are not assumed to be known. Instead, in the spirit of SDM (2004), we assume prior inclusion probabilities for the elements of $\mathbf{X}$ and $\mathbf{Z}$. In particular, we assume a prior inclusion probability of $\bar{n}/N$ for the variables in $\mathbf{X}$ and a prior inclusion probability of $\bar{m}/M$ for the variables in $\mathbf{Z}$, where $N = card(\mathbf{X})$ and $M = card(\mathbf{Z})$. This implies that the prior expected number of included $\mathbf{X}$-variables in the regression (excluding the constant) is $\bar{n}$ and the prior expected number of variables inducing nonlinearities is $\bar{m}$, leading to an expected model size of $(\bar{n}+1)(\bar{m}+1)$.

Given the choice of regressors from $\mathbf{X}$ and threshold variable from $\mathbf{Z}$ we proceed as follows to choose a threshold value z_j. We assign a diffuse prior to values of z_j actually observed in the sample after trimming $100\times\theta\%$ of the observations from each extreme of the empirical distribution. We impose this trimming of the distribution to avoid that one of the resulting regimes contains too few observations which could lead to unreliable estimation results. Therefore, the prior inclusion probability of $z_{j,i}$ (observation i in threshold variable z_j) as a threshold in (1) conditional on the inclusion of z_j as a threshold variable is uniform and given by $1/[T(1-2\theta)]$.[4]

It should be noted that this prior specification for the threshold values uses sample information and could thus be controversial if the Bayesian approach is to be taken literally. A related issue is the ordering of variables in the cross-sectional context, which is straightforward in the time-series context.[5] We proceed in the estimation by assuming a "natural ordering" of threshold variables $\mathbf{Z}$ from smallest to largest realized value and applying the trimming and selection of threshold to the ordered observations. Given the obvious difficulties involved in setting bounds to the prior distribution of the threshold values without observing the realized sample of the threshold variable and since using sample information for the prior specification is standard in the Bayesian literature of threshold estimation (see, for example, Koop and Potter, 1999), we decided to use this mixed approach to threshold estimation.[6]

Given the setting put forward above, the prior probability attached to a model containing n $\mathbf{X}$-variables and m threshold variables with thresholds $\{\tau_1, \ldots, \tau_m\}$ is[7]

$$P(M_{n,m,\tau_1,\ldots,\tau_m}) = (\bar{n}/N)^n(1-\bar{n}/N)^{N-n}(\bar{m}/M)^m(1-\bar{m}/M)^{M-m}[1/[T(1-2\theta)]]^m. \quad (2)$$

With this diffuse prior specification and further assuming a diffuse prior with respect to σ, the odds ratio for two models can be approximated (see Leamer, 1978, and Schwarz, 1978) as

[4] For two consecutive ordered observations $z_{j,i}$ and $z_{j,i+1}$, any threshold value τ_j in the interval $[z_{j,i}, z_{j,i+1})$ leads to the same variable $\mathsf{I}(z_j \leq \tau_j)$. This implies that we only need to define a discrete prior probability for each realized value of z_j, instead of a continuous prior density on the support of z_j.

[5] We thank Hashem Pesaran for raising this point in discussions with us.

[6] See also Phillips (1991) for a reference to this controversy.

[7] This is the prior model probability assuming that there are no repeated observations in the central 100(1-2θ)% of the empirical distribution of the variables in the $\mathbf{Z}$ group. If an observation for variable z_j repeated r times, its prior inclusion probability as a threshold value conditional on the inclusion of z_j as a threshold variable would be $r/[T(1-2\theta)]$, and $P(M_{n,m,\tau_1,\ldots,\tau_m})$ could be adjusted conveniently.

$$\frac{P(M_0|Y)}{P(M_1|Y)} = \frac{P(M_0)}{P(M_1)} T^{(k_0-k_1)/2} \left(\frac{SSE_0}{SSE_1}\right)^{-T/2}, \tag{3}$$

where k_i is the size of model i, $P(\cdot|Y)$ refers to posterior probabilities and SSE_i is the sum of squared residuals from the estimation of model i. Therefore, given our model space $\mathcal{M}$ the posterior probability of model i can be computed as

$$P(M_i|Y) = \frac{P(M_i)T^{-k_i/2}SSE_i^{-T/2}}{\sum_{j=1}^{card(\mathcal{M})} P(M_j)T^{-k_j/2}SSE_j^{-T/2}}. \tag{4}$$

The posterior model probabilities allow us to easily compute the first and second moment of the posterior densities of the α, β and τ parameters in (1), given by

$$E(\xi|Y) = \sum_{l=1}^{card(\mathcal{M})} P(M_l|Y)E(\xi|Y, M_l) \tag{5}$$

and

$$\begin{aligned} \text{var}(\xi|Y) \;=\; & \sum_{l=1}^{card(\mathcal{M})} P(M_l|Y)\text{var}(\xi|Y, M_l) + \\ & + \sum_{l=1}^{card(\mathcal{M})} P(M_l|Y)(E(\xi|Y, M_l) - E(\xi|Y))^2 \end{aligned} \tag{6}$$

where ξ is the parameter of interest and $E(\xi|Y, M_l)$ is the OLS estimator of ξ for the constellation of **X**- variables, **Z**-variables and threshold values implied by model l. The posterior probability that a given **X**-variable, **Z**-variable or threshold value is part of the regression can be computed as the sum of posterior model probabilities of those models containing the variable or threshold value of interest.

2.2 Random sampling in the BAT framework

Since the number of possible regressions for reasonable sizes of **X** and **Z** is enormous,[8] we have implemented both random sampling and a version of the "stratified" sampling procedure proposed by SDM (2004).[9] For the random sampler (RS), we use

[8] Notice that for a given **Z**-variable, $T(1-2\theta)$ threshold values are possible, and each threshold value defines a different model in our setting. This implies that, for a given group of **X** variables and two threshold variables, $[T(1-2\theta)]^2$ models are possible. For example, if T=90, θ=0.15, N=20 and M=2, $\mathcal{M}$ contains more than 4200 million models.

[9] For details see the Technical Appendix to SDM (2004), which is available at: `www.econ.cam.ac.uk/doppelhofer`

prior inclusion probabilities of variables in $\mathbf{X}$ and $\mathbf{Z}$ and the uniform prior over threshold values z_j to obtain (5), (6) and the posterior inclusion probabilities for $\mathbf{X}$-variables, $\mathbf{Z}$-variables and threshold values. The sampling design is as follows.

1. We sample n_j variables from $\mathbf{X}$ and m_j variables from $\mathbf{Z}$. Each variable in these sets has an inclusion probability of $\bar{n}/N$ and $\bar{m}/M$ for the set $\mathbf{X}$ and $\mathbf{Z}$ respectively.

2. For each one of the m_j $\mathbf{Z}$-variables sampled, we independently sample a threshold value from the empirical distribution of realized values after trimming $100 \times \theta\%$ of the observations from the extremes.

3. Equation (1) is estimated for the constellation of variables and threshold values which has been sampled. The information necessary in order to obtain equations (4), (5) and (6) are saved for the model sampled.

4. Steps 1.-3. are replicated R (a large number of) times and (4), (5) and (6) are computed using the replicated models, replacing $card(\mathcal{M})$ by R. Changes in parameters of interest are monitored to ensure convergence of averages of sampling distributions to the posterior distribuition[10].

The procedure allows us to obtain the posterior inclusion probability of all possible interactions of variables in $\mathbf{X}$ with indicator functions for a given variable of $\mathbf{Z}$ and a threshold value z_j. This posterior inclusion probability is computed as the sum of posterior model probabilities for models including that threshold variable and threshold value and allows us to obtain an estimate for the threshold value corresponding, for instance, to the mode of the posterior inclusion probability. Comparisons with the prior inclusion probabilities enable us to identify the threshold values whose inclusion probability increases or decreases after observing the data. In a similar fashion, the nonlinear effect can be evaluated by computing the posterior expected value and posterior variance of the parameter of the interaction for the corresponding threshold value.

2.3 Stratified sampling in the BAT framework

The "stratified sampler" first proposed by SDM (2004) extends naturally to the sampling over threshold variables $\mathbf{Z}$ and thresholds z_j proposed in this paper. The sampling design is very similar to the random sampling procedure described in 1-4

[10]See Doppelhofer and Durlauf (2006) for a discussion of model averaging techniques.

above. However, instead of the same (identical) prior sampling probability $\bar{n}/N$ and $\bar{m}/M$ for variables from $\mathbf{X}$ and $\mathbf{Z}$ respectively, we adjust the sampling probability to account for variables providing good fit. The first step is therefore replaced by:

1'. We sample n_j variables from $\mathbf{X}$ and m_j variables from $\mathbf{Z}$ starting with prior inclusion probability of $\bar{n}/N$ and $\bar{m}/M$, respectively. After a number of random draws, the sampling probabilities are adjusted to reflect model fit, captured by the *posterior inclusion probability* of variables in sets $\mathbf{X}$ and $\mathbf{Z}$,

$$P(\xi \neq 0|Y) = \sum_{l=1}^{S} P(M_l|Y)\mathsf{I}(\xi \neq 0|M_l)$$

for the corresponding parameter ξ. The sampling probabilities are then given by a weighted average of prior and posterior inclusion probabilities. To ensure that the posterior distributions are consistently estimated, the sampling probabilities are bounded away from the extremes (0 or 1) by introducing lower and upper bounds.

2'. Similarly, the probability of threshold z_j can be adjusted to reflect higher probability of choosing an important threshold point in the sampling probability. Inspecting the uniform prior and posterior distribution shows potential computational gains in simulating the posterior distribution.

Steps 3 and 4 are then repeated as in the random sampling procedure. The savings of computation time can be considerable, in particular in the case of a large number of nonlinear threshold interactions and threshold values.

3 Nonlinearities and growth: Empirical application

In this section we apply the BAT procedure to a reduced set of growth covariates in order to evaluate the existence and nature of nonlinearities in growth regressions. We choose the 21 variables that SDM (2004) find to be robustly related to growth using the (linear) BACE approach as the set $\mathbf{X}$. The variables are presented in a table in the Data Appendix, together with the rest of the variables included in the analysis carried out in SDM (2004).[11] For this application, we will use a relatively

[11] The first 18 variables are robustly related to growth meaning that, in the linear BACE setting, the posterior inclusion probability is higher than the prior inclusion probability. The other three

small group of variables as **Z**, formed by two variables that have often been reported to cause threshold-nonlinearity in growth regressions: the *initial level of GDP per capita* and the *proportion of years that the economy is open* according to the criteria in Sachs and Warner (1995). Durlauf and Johnson (1995), Hansen (2000), Masanjala and Papageorgiou (2004) and Crespo Cuaresma (2002) report evidence on nonlinearity induced by initial GDP per capita levels. Papageorgiou (2002) finds evidence that sets of countries with different openness levels tend to differ in the statistical model relating economic growth to other economic variables.[12]

The results presented below were obtained with ten million replications of the BAT procedure with random sampling setting $\bar{n}$=5, $\bar{m}$=1 and $\theta = 0.15$. We also ran the BAT procedure with other parameter constellations and the results concerning the existence and nature of nonlinearities appear robust to sensible changes in the expected number of included variables in the **X** group, $\bar{n}$, the expected number of included variables from the **Z** group, $\bar{m}$ and the trimming parameter.

Figure 1 and Figure 2 present the posterior inclusion probabilities for the threshold value in all possible interactions of the **X** group variables with each one of the threshold variable (initial GDP per capita in Figure 1 and proportion of years that the economy is open in Figure 2). The prior inclusion probability for each realized value is also plotted in the figures.[13] While in the case of initial GDP the prior inclusion probability is the same for all threshold values, in the case of the openness variable the repetition of identical values in the sample leads to different prior inclusion probabilities for each potential threshold value. The most remarkable feature of the posterior inclusion probabilities of the threshold values for initial GDP per capita is that they systematically fall below the prior inclusion probability, therefore lending little evidence to the existence of threshold nonlinearities caused by initial development levels once that model uncertainty is explicitly taken into account. The bigger bulk of posterior inclusion probability appears for many interactions in the interval

variables used as part of **X** (DENS60, RERD and OTHFRAC) are marginally related to growth: the posterior inclusion probability is slightly smaller than the prior inclusion probabilities, but their corresponding effect is estimated with high precision when they are included in the growth regression.

[12]See also Huang and Chang (2006) and Papageorgiou (2006).

[13]For a given interaction and a threshold value, the prior inclusion probability is given by the product of the prior inclusion variable of the corresponding **X** variable ($\bar{n}/N$), the corresponding **Z** variable ($\bar{m}/M$) and the corresponding threshold value ($r/[T(1-2\theta)]$, where r is the number of times the threshold value is repeated in the range of potential threshold values of the **Z** variable).

Figure 1: Posterior and prior inclusion probability, threshold value in *Initial GDP per capita*

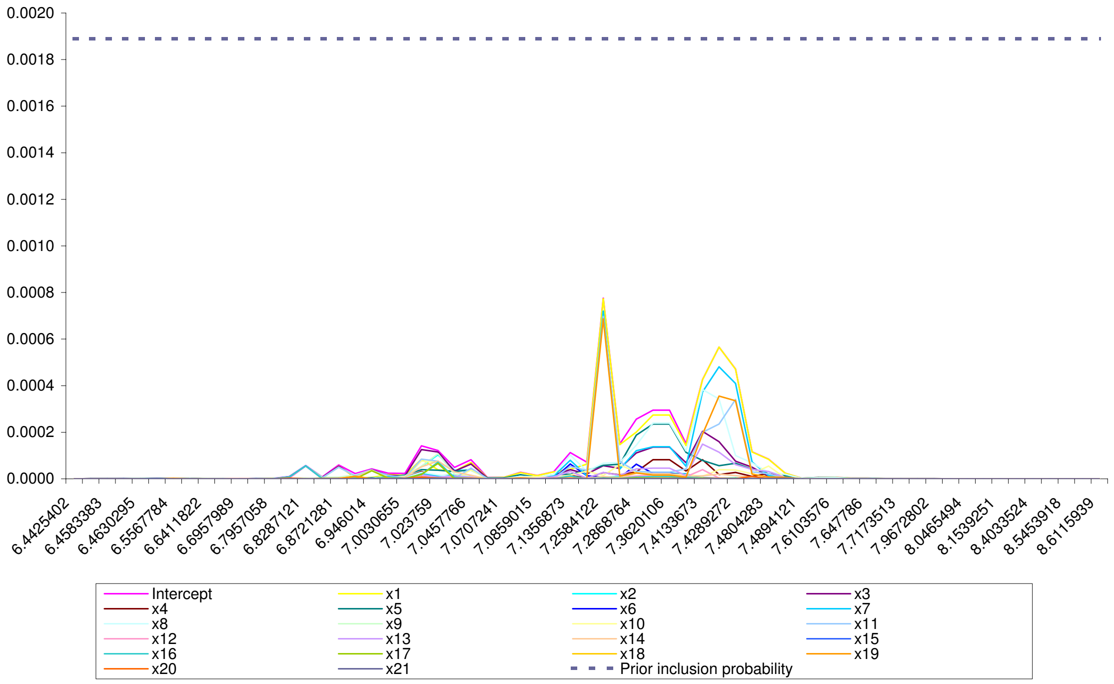
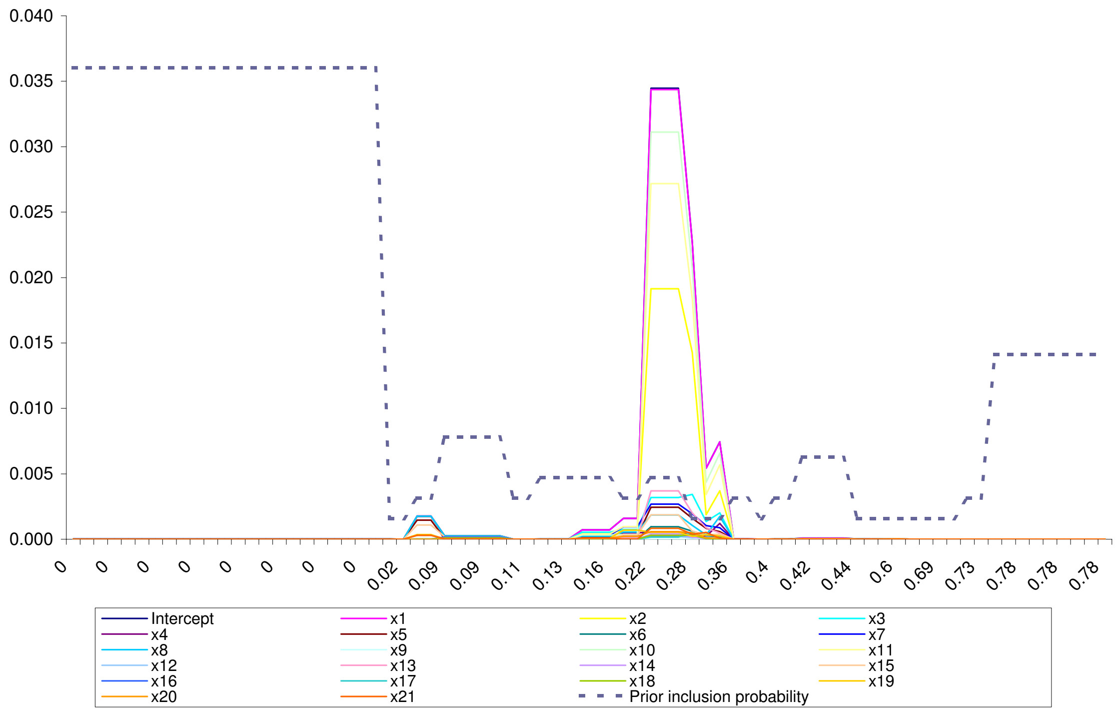


Figure 2: Posterior and prior inclusion probability, threshold value in proportion of *Years Open*



between 7.26 (corresponding to the initial GDP per capita of Malaysia) and 7.45 (corresponding to the initial GDP per capita of Algeria). It should be noted that for simulations run setting $\bar{m}$=2 (that is, considering only nonlinear threshold models with both threshold variables as the relevant class), posterior inclusion probabilities in this range appeared greater than the prior inclusion probabilities, but as long as model uncertainty with respect the existence of nonlinearities is taken into account (that is, for parameter constellations with $\bar{m} < 2$ such as the one reported here), the evidence of threshold effects caused by initial GDP per capita levels disappears.

Variable	β: Posterior mean	β: Posterior s.d.	β^*: Posterior mean	β^*: Posterior s.d.
Intercept	0.060352	0.022257	-0.009038	0.014481
East Asian dummy	0.019399	0.006475	-0.038349	0.010120
Primary schooling 1960	0.025717	0.010226	0.017526	0.015585
Investment price	-0.000083	0.000027	0.000011	0.000085
Fraction tropical area	-0.013797	0.004492	0.008528	0.008710
Malaria prevalence	-0.011639	0.008979	-0.018294	0.019087
Life expectancy 1960	0.000708	0.000351	-0.000719	0.000653
African dummy	-0.008482	0.011845	-0.031641	0.009293
Latin American dummy	-0.012638	0.005627	0.008361	0.007484
Spanish colony	-0.009723	0.005534	0.010712	0.007640

Values obtained with ten million replications of the BAT procedure for the group of robust variables in SDM (2004) (first 21 variables in the Data Appendix), for $\bar{n}$=5, $\bar{m}$=1 and $\theta = 0.15$. Posterior mean and standard deviation of β^* evaluated at the threshold value of openness corresponding to the mode of the posterior inclusion probability of each interaction reported.

Table 1: Posterior mean and standard deviations of β and β^* conditional on inclusion for openness as a threshold variable

In Figure 2 the posterior inclusion probabilities for the threshold value corresponding to the openness variable are presented. In the case of this threshold variable posterior inclusion probabilities are higher than prior inclusion probabilities in the range delimited by 0.22 (corresponding to the openness experience of Gambia and Ghana in our sample) and 0.33 (the proportion of years open for Nicaragua and Syria in our data) for the interactions with the following variables: the regression intercept, East Asian dummy, primary schooling 1960, investment price, fraction of tropical area, malaria prevalence, life expectancy in 1960, African dummy, Latin American dummy and Spanish colony. For these variables, Table 2 presents the posterior mean and standard deviation of β and β^* in (1) conditional on inclusion of the respective variables, evaluated at the threshold value of the openness variable corresponding to the mode of the posterior inclusion probability for each interaction.

The interaction effect is very well estimated for the case of the East Asian and African dummies, and the results shed an interesting light on the effects which are picked up by these variables in cross country growth regressions. The posterior mean of the East Asian dummy parameter (conditional on inclusion) corresponding to the regime of "open countries" (defined by a threshold parameter of 0.22 in the variable "Years open", which corresponds to the mode of the posterior inclusion parameter) is very similar to the result obtained in SDM (2004)[14] for the linear setting and is estimated very precisely. The posterior mean of the additive effect for observations in the regime of "closed countries" is -0.038, with a posterior standard deviation of 0.010, which deems the positive effect of the East Asian dummy inexistent for this subsample. A similar conclusion is reached for the case of the African dummy: when the interaction effects with openness are taken into account, this variable appears only robust and estimated with a high degree of precision in the regime corresponding to the subsample of relatively closed countries. Furthermore, the quantitative effect in this regime is estimated to be higher in absolute value than the linear elasticity obtained in SDM (2004).[15] These results suggest that these regional dummies are basically picking up the effect of subsamples of countries with a differential openness experience in the period under consideration.

4 Conclusions

We propose a new method of jointly assessing threshold effects and model uncertainty in the framework of cross-country growth regressions. Our methodology makes use of Bayesian model averaging in the spirit of SDM (2004) to deal with model uncertainty, including uncertainty about nonlinear effects. We put forward a method for estimation of thresholds based on the evaluation of the posterior inclusion probability of potential threshold values.

[14]In SDM (2004)'s results, the East Asian dummy is found to be the most robust variable of a set of 67 growth covariates. Conditional on inclusion of this variable in the linear setting, the posterior mean of the parameter attached to the dummy in SDM (2004) is 0.022, with posterior standard deviation of 0.006.

[15]The posterior mean conditional on inclusion for the African dummy in SDM (2004) is -0.015, with a posterior standard deviation of 0.007.

We use the set of explanatory variables that SDM (2004) found to be robustly related to economic growth in linear models. As threshold variables, we use initial GDP per capita and the proportion of years that the economy was open. We find no evidence of nonlinear growth effects generated by initial level of GDP per capita. This is contrary to other empirical studies (see for instance Durlauf and Johnson, 1995, and Hansen, 2000) which do not explicitly take model uncertainty into account, whereas we allow for uncertainty about model size, threshold values and the nature of the interactions. We find evidence of robust interactions between the number of years an economy has been open and several other growth determinants. Our results imply that the widely used East Asian dummy and African dummy are picking up the effect of subsamples of countries with a high and low degree of openness, respectively.

References

[1] Azariadis, C. and A. Drazen (1990), Threshold externalities in economic development. *Quarterly Journal of Economics* 105, 501-526.

[2] Barro, R.J. (1991), Economic growth in a cross section of countries. *Quarterly Journal of Economics* CVI, 407-443.

[3] Brock, W.A. and S.N. Durlauf (2001), Growth Empirics and Reality. *The World Bank Economic Review* 15, 229-72.

[4] Crespo Cuaresma, J. (2002), Some million thresholds: nonlinearity and cross-country growth regressions, *Working Paper 0210*, Department of Economics, University of Vienna.

[5] Doppelhofer, G. and S.N. Durlauf (2006), Model Averaging: A Survey. *Mimeo.*

[6] Durlauf, S.N. and P.A. Johnson (1995), Multiple regimes and cross-country growth regressions. *Journal of Applied Econometrics* 10, 365-384.

[7] Durlauf, S.N., Kourtellos, A. and A. Minkin (2001), The local Solow model. *European Economic Review* 45, 928-940.

[8] Durlauf, S. N. and D.T. Quah (1999), The new empirics of economic growth, in J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, vol. IA. North-Holland, Amsterdam, 231-304.

[9] Fernández, C., Ley, E. and M.F. Steel (2001), Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics* 16, 563-576.

[10] Hansen, B. E. (1996), Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 64, 413-430.

[11] Hansen, B. E. (2000), Sample splitting and threshold estimation. *Econometrica*, 68, 575-603.

[12] Huang, H. and Y. Chang (2006), Trade as a threshold variable for multiple regimes: A comment. *Economics Letters* 91, 458-459.

[13] Koop, G. and S. M. Potter (1995), Bayes factors and nonlinearity: Evidence from economic time series. *Journal of Econometrics* 88, 251-281.

[14] Kormendi, R., and P. Meguire (1985), Macroeconomic determinants of growth, cross-coutry evidence. *Journal of Monetary Economics* 16, 141-163.

[15] Leamer, E.E. (1983), Let's take the Con out of Econometrics. *American Economic Review* 73, 31-43.

[16] Levine, R. and D. Renelt (1992), A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82, 942-963.

[17] Masanjala, W.H. and C. Papageorgiou (2004), The Solow Model with CES Technology: Nonlinearities and parameter heterogeneity. Journal of Applied Econometrics 19, 171-201.

[18] Papageorgiou, C. (2002), Trade as a threshold variable for multiple regimes. *Economics Letters* 77, 85-91.

[19] Papageorgiou, C. (2006), Trade as a threshold variable for multiple regimes: Reply. *Economics Letters* 91, 460-461.

[20] Phillips, P.C.B. (1991), To criticize the critics: An objective Bayesian analysis of stochastic trends. *Journal of Applied Econometrics* 6, 333-364.

[21] Sachs, J. and A. Warner (1995), Economic reform and the process of global integration. *Brookings Papers on Economic Activity* 1, 1-95.

[22] Sala-i-Martin, X. (1997a), I just ran two million regressions. *American Economic Review* 87, 178-183.

[23] Sala-i-Martin, X. (1997b), I just ran four million regressions. *NBER Working Paper 6252*.

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

A Data Appendix

Rank	Short Name	Variable Description	PIP	Mean	S.D.
Dep. Var.	GROWTH	Growth of GDP per capita at PPP between 1960–1996.	–	0.0182	0.0191
1	EAST	East Asian Dummy	0.82	0.11364	0.31919
2	P60	Primary Schooling Enrollment	0.80	0.72614	0.29321
3	IPRICE1	Investment Price	0.77	92.47	53.68
4	GDPCH60L	Log GDP in 1960	0.68	7.35494	0.90108
5	TROPICAR	Fraction of Tropical Area	0.56	0.57024	0.47160
6	DENS65C	Population Coastal Density	0.43	146.87	509.83
7	MALFAL66	Malaria Prevalence	0.25	0.33943	0.43089
8	LIFE060	Life Expectancy	0.21	53.72	12.06
9	CONFUC	Fraction Confucian	0.21	0.01557	0.07932
10	SAFRICA	Sub-Saharan Africa Dummy	0.15	0.30682	0.46382
11	LAAM	Latin American Dummy	0.15	0.22727	0.42147
12	MINING	Fraction GDP in Mining	0.12	0.05068	0.07694
13	SPAIN	Spanish Colony Dummy	0.12	0.17045	0.37819
14	YRSOPEN	Years Open 1950-94	0.12	0.35545	0.34445
15	MUSLIM00	Fraction Muslim	0.11	0.14935	0.29616
16	BUDDHA	Fraction Buddhist	0.11	0.04659	0.16760
17	AVELF	Ethnolinguistic Fractionalization	0.10	0.34761	0.30163
18	GVR61	Gov't Consumption Share	0.10	0.11610	0.07454
19	DENS60	Population Density	0.09	108.07	201.44
20	RERD	Real Exchange Rate Distortions	0.08	125.03	41.71
21	OTHFRAC	Fraction Speaking Foreign Language	0.08	0.32092	0.41363

Explanatory variables are ranked by Posterior Inclusion Probability $P(\xi_j \neq 0|Y)$ (PIP) using the BACE method (SDM, 2004). The set of regressors **X** is given by variables 1 to 21. The threshold variables **Z** are ranked 4 (Log GDP in 1960) and 14 (Years Open 1950-94), respectively, but this is not necessarily informative of their role as threshold variable. Variables ranked 22 to 67 were not included in the results presented.

Rank	Short Name	Variable Description	PIP	Mean	S.D.
22	**OPENDEC1**	**Openness Measure 1965-74**	0.08	0.52307	0.33591
23	**PRIGHTS**	**Political Rights**	0.07	3.82250	1.99661
24	**GOVSH61**	**Government Share of GDP**	0.06	0.16636	0.07115
25	**H60**	**Higher Education Enrollment**	0.06	0.03761	0.05006
26	**TROPPOP**	**Fraction Population In Tropics**	0.06	0.29998	0.37311
27	**PRIEXP70**	**Primary Exports**	0.05	0.71988	0.28270
28	**GGCFD3**	**Public Investment Share**	0.05	0.05216	0.03882
29	**PROT00**	**Fraction Protestant**	0.05	0.13540	0.28506
30	**HINDU00**	**Fraction Hindu**	0.04	0.02794	0.12465
31	**POP1560**	**Fraction Population Less than 15**	0.04	0.39251	0.07488
32	**AIRDIST**	**Air Distance to Big Cities**	0.04	4324	2614
33	**GOVNOM1**	**Nominal Government Share**	0.04	0.14898	0.05843
34	**ABSLATIT**	**Absolute Latitude**	0.03	23.21	16.84
35	**CATH00**	**Fraction Catholic**	0.03	0.32826	0.41459
36	**FERTLDC1**	**Fertility**	0.03	1.56202	0.41928
37	**EUROPE**	**European Dummy**	0.03	0.21591	0.41381
38	**SCOUT**	**Outward Orientation**	0.03	0.39773	0.49223
39	**COLONY**	**Colony Dummy**	0.03	0.75000	0.43549
40	**CIV72**	**Civil Liberties**	0.03	0.50947	0.32593
41	**REVCOUP**	**Revolutions and Coups**	0.03	0.18489	0.23223
42	**BRIT**	**British Colony Dummy**	0.03	0.31818	0.46844
43	**LHCPC**	**Hydrocarbon Deposits**	0.02	0.42115	4.35121
44	**POP6560**	**Fraction Population Over 65**	0.02	0.04881	0.02898
45	**GDE1**	**Defense Spending Share**	0.02	0.02589	0.02463
46	**POP60**	**Population in 1960**	0.02	20308	52538
47	**TOT1DEC1**	**Terms of Trade Growth in 1960s**	0.02	-0.00208	0.03455
48	**GEEREC1**	**Public Education Spending Share**	0.02	0.02441	0.00964
49	**LANDLOCK**	**Landlocked Country Dummy**	0.02	0.17045	0.37819
50	**HERF00**	**Religion Measure**	0.02	0.78032	0.19321
51	**SIZE60**	**Size of Economy**	0.02	16.15	1.82
52	**SOCIALIST**	**Socialist Dummy**	0.02	0.06818	0.25350
53	**ENGFRAC**	**English Speaking Population**	0.02	0.08398	0.25224
54	**PI6090**	**Average Inflation 1960-90**	0.02	13.13	14.99
55	**OIL**	**Oil Producing Country Dummy**	0.02	0.05682	0.23282
56	**DPOP6090**	**Population Growth Rate 1960-90**	0.02	0.02153	0.00946
57	**NEWSTATE**	**Timing of Independence**	0.02	1.01136	0.97667
58	**LT100CR**	**Land Area Near Navigable Water**	0.02	0.47216	0.38021
59	**SQPI6090**	**Square of Inflation 1960-90**	0.02	394.54	1119.70
60	**WARTIME**	**Fraction Spent in War 1960-90**	0.02	0.06955	0.15241
61	**LANDAREA**	**Land Area**	0.02	867189	1814688
62	**ZTROPICS**	**Tropical Climate Zone**	0.02	0.19002	0.26869
63	**TOTIND**	**Terms of Trade Ranking**	0.02	0.28127	0.19038
64	**ECORG**	**Capitalism**	0.02	3.46591	1.38089
65	**ORTH00**	**Fraction Orthodox**	0.02	0.01867	0.09829
66	**WARTORN**	**War Participation 1960-90**	0.02	0.39773	0.49223
67	**DENS65I**	**Interior Density**	0.02	43.37	88.06