

TI 2011-012/4
Tinbergen Institute Discussion Paper



Variable Selection and Functional Form Uncertainty in Cross-Country Growth Regressions

Tim Salimans

Erasmus University Rotterdam, and Tinbergen Institute.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Variable selection and functional form uncertainty in cross-country growth regressions

Tim Salimans
Tinbergen Institute
Erasmus University Rotterdam

Abstract

Regression analyses of cross-country economic growth data are complicated by two main forms of model uncertainty: the uncertainty in selecting explanatory variables and the uncertainty in specifying the functional form of the regression function. Most discussions in the literature address these problems independently, yet a joint treatment is essential. We perform this joint treatment by extending the linear model to allow for multiple-regime parameter heterogeneity of the type suggested by new growth theory, while addressing the variable selection problem by means of Bayesian model averaging. Controlling for variable selection uncertainty, we confirm the evidence in favor of new growth theory presented in several earlier studies. However, controlling for functional form uncertainty, we find that the effects of many of the explanatory variables identified in the literature are not robust across countries and variable selections.

Keywords: growth regression, variable selection, model uncertainty, model averaging, semi-parametric Bayes, MCMC.

JEL codes: C11, C14, C15, O40, O57.

Acknowledgements: Special thanks go to my supervisor Richard Paap, whose advice was a great help in realizing this paper. In addition, I thank all those present at the ESOBE conference 2010 for their insightful comments and remarks on this work. In particular, I am grateful to Mark Steel for suggesting a more explicit nesting of the linear model specification in the model presented here. Finally, I thank The Netherlands Organization for Scientific Research for financially supporting this project.

Address for correspondence: Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000DR, Rotterdam, The Netherlands. E-mail: salimans@ese.eur.nl

1 Introduction

Many economic studies aim to determine the driving factors of economic growth. Following the seminal work of Kormendi and Meguire (1985) and Barro (1991), an important tool in this endeavor has been the cross-country growth regression, i.e. the use of regression analysis to determine what variables are correlated with economic growth in a cross-section of countries. The literature has identified two major problems with this technique. The first is that there is only a limited number of countries and a potentially very large number of variables to explain economic growth. The decision of which variables to include in the regression therefore has a strong influence on the conclusions that are drawn from the analysis. Since this decision is often guided by nothing but the whim of the researcher, there is no guarantee that these conclusions are not the product of data mining and selective presentation of data (see Leamer (1983) and Geweke (2005, sections 8.4 and 8.5)).

The second objection raised against cross-country growth regression is that most studies unreasonably restrict attention to the set of linear regression models. The linear regression model complies with the classical Solow model (see Mankiw et al., 1992) which specifies that log output is an additive linear function of technology, capital and labor. However, a range of new growth models, collectively known as *new growth theory*, pose the existence of multiple steady states in economic growth (see Aghion et al., 1999). Although these models typically specify the growth path of each country to be linear in its variables, the slopes of the growth path depend on which steady state the country is in. Since the steady state of a country depends on its initial conditions, such as its level of economic development and human capital, the process determining economic growth in these models is nonlinear in the regressors.

Both issues have received much discussion in the literature, however only rarely in the same paper. Yet, a joint treatment of these two sources of model uncertainty is absolutely essential: Variable selection methods do not necessarily select the same variables under different model specifications, and evidence of nonlinearity may not hold up under different variable selections. To examine these issues, this paper presents an integrated analysis of variable selection and functional form specification in cross-country growth regressions. We perform this joint treatment by extending the linear growth regression model to explicitly allow for multiple-regime parameter heterogeneity as suggested by new growth theory, while simultaneously addressing the variable selection problem by performing a Bayesian model averaging. Estimating the new models on the data set of Sala-i-Martin et al. (2004) provides evidence of multiple-regime parameter heterogeneity of the type predicted by new growth theory and empirically documented by Durlauf and Johnson (1995) and Liu and Stengos (1999). In addition, we find that many of the explanatory variables indicated by the literature do not have robust marginal effects across countries when allowing for a more flexible model specification, contradicting the results of Minier (2007). Our results offer some new insights into the form of the parameter heterogeneity in the growth data and we discuss its connection to phenomena like the natural resources curse.

The outline of the remainder of this paper is as follows. Section 2 provides a short review of the existing literature on growth regression. The statistical methodology of the paper is explained in Section 3, where we introduce a new set of models that allows for multiple-regime nonlinearity. In Section 4 we present the estimation results for these models and compare them with the linear model specification. Finally, Section 5 concludes.

2 Robustness in growth regressions

The large body of literature on cross-country growth regression started with the work of Kormendi and Meguire (1985), Grier and Tullock (1989) and Barro (1991). Since then the literature has identified a large number of variables correlated to economic growth. However, these variables were not discovered by the analysis of an ever greater amount of data, but rather by the specification of an ever greater amount of different models, casting doubt onto the statistical validity of these findings. Attention to this problem was first raised by the influential paper of Levine and Renelt (1992), who investigated the robustness of earlier findings to different model specifications by employing a variant of the extreme bounds analysis of Leamer (1983). This analysis proceeds by estimating the coefficient of a regressor in many different linear regression models, each controlling for a different subset of regressors, and analyzing the different results. If the regressor of interest is found to be significantly different from zero in each regression, with the same sign, the influence of the regressor is called *robust*. Otherwise it is called *fragile*. Levine and Renelt (1992) (henceforth LR) found that many of the relationships uncovered in earlier work on growth were in fact fragile.

The paper by LR was followed by many responses from the growth community. Although widely appreciated for bringing attention to the explosion of different model specifications in the growth literature, its methodology received criticism from several authors. An influential response was the one by Sala-i-Martin (1997), who argued that the extreme bounds approach was overly stringent. An important property of the approach is that a negative result from a single model specification can potentially negate the positive results from a much larger number of models, even if those other models fit the data much better. Sala-i-Martin argued that this property, combined with the large number of different model specifications, was almost guaranteed to produce the negative results reported by LR. As an alternative method of providing robust inference, he proposed to instead look at the average result of the regressions, with each model receiving a weight proportional to its data likelihood. His follow-up paper (Sala-i-Martin et al., 2004) further developed this approach by deriving a new weighting method based on Bayesian model averaging. Another contribution using the concept of model averaging is Fernandez et al. (2001b) who presented a similar analysis.

A second criticism the LR study received is that it unreasonably restricted attention to the set of linear models, while new growth theory predicts a nonlinear relationship between growth and the explanatory variables. This criticism was backed up by empirical evidence

provided by Durlauf and Johnson (1995), among many others, who documented the existence of multiple-regime parameter heterogeneity. By performing a tree regression, they allocated the countries in their data set into multiple regimes based on initial conditions related to the levels of economic development and human capital of the country. Liu and Stengos (1999) confirmed these results by estimating a classical semi-parametric model on the LR dataset, modeling the same kind of multiple-regime nonlinearity. They found that their nonlinear model improved upon the linear specification. Additional evidence in support of heterogeneity was found by Paap et al. (2005) and Basturk et al. (2010) who modeled economic growth using mixtures of linear regression models. While these studies took into account the uncertainty in the functional form of the growth equation, they considered only small fixed sets of explanatory variables, ignoring the uncertainty in the variable selection process.

To the author's knowledge, Minier (2007) and Cuaresma and Doppelhofer (2007) are the only attempts to date at combining both sources of uncertainty in an analysis. Minier (2007) investigated the influence of nonlinearities in the fragile variables of LR by repeating their analysis, but subsequently introducing quadratic and interaction terms into the specification as well as allowing for different growth regimes by splitting the sample according to initial conditions. By doing so, several more variables relating to fiscal policy appeared robust in her specification compared to the original LR analysis. However, in determining the robustness of the variables she only looked at the parameters of the linear terms, ignoring the coefficients of the higher order regressors. This ignores the fact that, through the quadratic and interaction terms, the 'robust' regressors may very well have marginal effects with different signs in different specifications, which makes her conclusions difficult to compare with the original LR results.

Cuaresma and Doppelhofer (2007) extended the approach of Sala-i-Martin et al. (2004) to allow for threshold effects in the model specification. Similar to the analysis of Durlauf and Johnson (1995), they effectively split the sample based on explanatory variables, but instead of defining the splitting thresholds a priori they estimated them together with the regression coefficients. Using this approach, they performed Bayesian model averaging over the subset of variables found to be robustly correlated to growth by Sala-i-Martin et al. (2004). Contrary to the results of Durlauf and Johnson (1995), Liu and Stengos (1999) and others they did not find much evidence for nonlinearity. It is worth investigating whether this is due to their particular model specification or their (limited) consideration of variable selection uncertainty. In addition, a further investigation into the influence of possible nonlinearities on the variable selection problem is needed.

3 Statistical Methodology

Following the tradition of the growth literature, as exemplified by Barro (1991), Levine and Renelt (1992), Sala-i-Martin et al. (2004) and many others, we will consider linear regression models of the form

$$y_i = \alpha + \beta_{i,1}x_{i,1} + \beta_{i,2}x_{i,2} + \dots + \beta_{i,p}x_{i,p} + \epsilon_i, \quad \epsilon_i \sim NID(0, \sigma^2) \quad (1)$$

where y_i is the i -th country's long term growth rate, $\{x_{i,1}, \dots, x_{i,p}\}$ is a collection of p explanatory variables, α is an intercept, $\{\beta_{i,1}, \dots, \beta_{i,p}\}$ is a collection of (possibly country-specific) regression coefficients and ϵ_i is an error term. Most recent work on growth regressions has focused on the selection of explanatory variables to include in this model (Fernandez et al. (2001b), Sala-i-Martin et al. (2004)) or on the specification of country-specific parameters (Kalaitzidakis et al. (2000), Minier (2007), Paap et al. (2005), Basturk et al. (2010), Maasoumi et al. (2007)). A rigorous regression analysis of cross-country growth data should simultaneously take into account both of these sources of uncertainty. Bayesian analysis is ideally suited for this task as it offers a systematic method of quantifying this uncertainty that is not offered by classical statistics. Indeed, the majority of the literature dealing with these issues in growth regressions builds on the Bayesian framework (e.g. Levine and Renelt (1992), Sala-i-Martin (1997), Fernandez et al. (2001b), Sala-i-Martin et al. (2004), Minier (2007) and many others).

3.1 Prior specification for the regression parameters

The uncertainty due to the unknown β parameters can be quantified in terms of a prior distribution. Since the model (1) contains more parameters than we have observations, the specification of this prior distribution will have an important effect on our analysis. The earlier studies by Sala-i-Martin (1997), Fernandez et al. (2001b) and Sala-i-Martin et al. (2004) all assume that the model parameters are the same across countries. Subject to this constraint, they use the popular conjugate g-prior introduced by Zellner (1986) to complete the specification. The advantage of this prior is its analytical tractability and its invariance to the scale of the regressors. Stacking the $\{\beta_{i,1}, \dots, \beta_{i,p}\}$ parameters into a vector β , this prior is given by

$$\beta|X, \sigma^2 \sim N[0, \sigma^2(gX'X)^{-1}] \quad (2)$$

where X is the matrix of explanatory variables included in the model and g is a parameter to be set by the researcher. Sala-i-Martin et al. (2004) set $g = 1/n$ and then take an approximation based on the sample size n becoming large. Alternatively, Fernandez et al. (2001b) set the g parameter to $1/k^2$, where k is the number of candidate regressors. They choose this constant because they find it generally leads to good estimation results, as they show in Fernandez et al. (2001a). We choose $g = 1/n$ as this choice has the intuitively attractive property of keeping the scale of the prior variance constant when the sample size changes.

The assumption that all countries have the same parameters is consistent with the Solow model, as discussed in the introduction, but not with new growth theory. According to

new growth theory, countries end up in different growth regimes when they are subject to different initial conditions. These different regimes create nonlinearities in the growth data, as documented by Durlauf and Johnson (1995), Liu and Stengos (1999), Kalaitzidakis et al. (2000), Minier (2007) and others. A prior specification that puts 100% probability on the parameters being equal across countries seems unreasonably dogmatic as it completely rules out such nonlinearities a priori. Here we present one way of allocating prior probability to new growth theory. To facilitate comparison we start out by adopting (2) as the marginal prior distribution for the country specific parameters β_i as in the earlier studies. However, instead of making the classical assumption that the parameters are equal across countries, we follow new growth theory in allowing them to vary according to initial conditions. This is the intuition behind the work of Durlauf and Johnson (1995) and Minier (2007), who split the sample according to initial output and human capital and find evidence of parameter heterogeneity between the different groups. A problem with this approach is that it is not clear a priori at what level the sample should be split or how many of these splits should be made. This makes it hard to do derive statistical conclusions from such a procedure, as discussed by Durlauf and Johnson (1995). Instead, we formalize the intuition of new growth theory by introducing prior covariances between the country-specific parameter vectors

$$\text{Cov}(\beta_i, \beta_j) = \rho_{i,j} \sigma^2 (gX'X)^{-1} \quad (3)$$

where $\rho_{i,j}$ is the prior correlation between the parameters for country i and country j . In other words, we specify a joint multivariate normal prior for all β parameters.

$$\beta = (\beta'_1, \dots, \beta'_n)' \sim N(0, \rho \otimes \sigma^2 (gX'X)^{-1}) \quad (4)$$

where ρ is the matrix of all cross-country correlations $\rho_{i,j}$. New growth theory suggests that countries with similar initial conditions are likely to have similar parameters. Of these initial conditions, output and human capital are believed to be the most important (e.g. Durlauf and Johnson (1995), Liu and Stengos (1999), Minier (2007)). We formalize this idea by specifying $\rho_{i,j}$ as a Gaussian distance function containing measures for output and human capital.

$$\rho_{i,j} = \exp(-\gamma[(O_i - O_j)^2/\text{Var}(O) + (H_i - H_j)^2/\text{Var}(H)]) \quad (5)$$

where O denotes log initial GDP per capita and H denotes initial human capital as captured by the level of higher education enrollment, although using primary schooling instead turns out to give very similar results. Conditional on the α , γ and σ^2 parameters, our prior for economic growth as a function of the explanatory variables now corresponds to the following Gaussian process.

$$\begin{aligned}
y(x)|X, \alpha, \gamma, \sigma^2 &\sim \mathcal{GP}[\alpha, K(x_i, x_j) + \sigma^2 \delta_{i,j}] \\
K(x_i, x_j) &= \text{Cov}(y_i, y_j | x_i, x_j) = \sigma^2 \rho_{i,j} x_i (gX'X)^{-1} x_j'
\end{aligned} \tag{6}$$

A Gaussian process is a stochastic process of which any finite dimensional distribution is distributed according to a (multivariate) Gaussian. Our conditional prior distribution for the vector of growth rates corresponding to the countries in the sample is thus the following multivariate normal distribution

$$\begin{aligned}
y|X, \alpha, \gamma, \sigma^2 &\sim N(\alpha \iota, \Sigma) \\
\Sigma_{i,i} &= K(x_i, x_i) + \sigma^2, \quad \Sigma_{i,j} = K(x_i, x_j)
\end{aligned} \tag{7}$$

where ι is a vector of ones. Gaussian processes can be used to specify very flexible prior specifications that allow for nonlinearities while remaining analytically tractable. For this reason they are becoming increasingly popular, especially in the field of machine learning, see e.g. Rasmussen and Williams (2006). By adapting the covariance function $K(x_i, x_i)$, the approach presented here can easily be modified to allow for almost any form of non-linearity. One such way of specifying the covariance function is the method of Koop and Poirier (2004), who explicitly discuss the specification of variants of commonly used classical semi-parametric methods in this form.

The prior cross-country correlations $\rho_{i,j}$ are determined by the γ parameter through equation (5). This parameter is somewhat difficult to interpret directly, but it has a simple one-to-one relationship with the median of the correlations $\rho_{i,j}$.

$$\bar{\rho} = \text{median}(\rho_{i,j}) = \exp(-\gamma \bar{d}) \tag{8}$$

where \bar{d} is the median of the 'distances' $(O_i - O_j)^2 / \text{Var}(O) + (H_i - H_j)^2 / \text{Var}(H)$ in equation (5). If the median prior correlation $\bar{\rho}$ is set to one, then γ is equal to zero, all β_i parameters are assumed to be equal and our model reduces to the linear model specification of Fernandez et al. (2001b) and Sala-i-Martin et al. (2004). If on the other hand $\bar{\rho}$ is set to zero, γ is infinite and the regression coefficients of the different countries are assumed to be completely independent. We would like to infer the right amount of dependence from the data and we therefore assume a uniform prior on $\bar{\rho}$, which implies an exponential prior on γ .

$$\bar{\rho} \sim U[0, 1] \quad \Leftrightarrow \quad p(\gamma) = \bar{d} \exp(-\bar{d}\gamma) \tag{9}$$

Finally, we finish the prior specification by adopting standard non-informative priors for the remaining parameters α and σ^2 as is often suggested in the literature on Bayesian model averaging (e.g. Fernandez et al., 2001a).

$$\begin{aligned}
p(\sigma^2) &\propto 1/\sigma^2 \\
p(\alpha) &\propto 1
\end{aligned} \tag{10}$$

Conditional on X and $\bar{\rho}$, the proportional marginal data likelihood and posterior distributions can now be obtained analytically by treating the uninformative prior on α as the limit of a normal distribution, i.e. $\alpha \sim N(0, \delta)$ with $\delta \rightarrow \infty$. Applying the matrix determinant lemma and Sherman-Morrison formula to the covariance matrix of y then leads to the following expression for the marginal likelihood.

$$p(y|X, \bar{\rho}) \propto \frac{1}{\sqrt{|\iota' B^{-1} \iota|}} |B|^{-1/2} \left(y' \left[B^{-1} - \frac{B^{-1} \iota \iota' B^{-1}}{\iota' B^{-1} \iota} \right] y \right)^{-(n-1)/2} \tag{11}$$

where $B = \Sigma/\sigma^2$ is proportional to the covariance matrix and n is the number of observations in y . A similar derivation shows that the posterior distribution for σ^2 is inverted Gamma and that for β multivariate Student's t .

$$\begin{aligned}
\sigma^2|y, X, \bar{\rho} &\sim \text{IG}(\nu/2, \theta/2), \quad \nu = n - 1, \quad \theta = y' B^{-1} y - (y' B^{-1} \iota)^2 / (\iota' B^{-1} \iota) \\
\beta|y, X, \bar{\rho} &\sim t_{np}(\nu, \mu, S), \quad \mu = C B^{-1} y, \quad S = \frac{\theta}{\nu} (\rho \otimes (gX'X)^{-1} - C B^{-1} C') \tag{12}
\end{aligned}$$

where p is the number of explanatory variables in the model, and C is proportional to the prior covariance between β and y , characterized by $C_{i+pm,j} = \rho_{m+1,j} [(gX'X)^{-1}_{i,1} X_{j,1} + \dots + (gX'X)^{-1}_{i,p} X_{j,p}]$ for integers $i \leq p, m < n, j \leq n$.

3.2 Variable selection

A second source of uncertainty is the problem of selecting regressors to include in the model. The data set we consider contains 88 countries and a list of 67 potential explanatory variables, giving 2^{67} different possible subsets of regressors to include in our model. All of these variables can reasonably be expected to influence economic growth and we cannot be sure a priori which subset of variables we should use. In such cases we can proceed by assigning a prior probability $P(m_j)$ to each of these different models, representing our assessment of the probability of that model being correct before having seen the data. Given the data $y = \{y_1, \dots, y_n\}$, we can then obtain posterior model probabilities by applying Bayes' rule.

$$P(m_j|y) = \frac{P(y|m_j) P(m_j)}{\sum_l P(y|m_l) P(m_l)} \tag{13}$$

where $P(y|m_j)$ is the likelihood of the data under model j . Our final conclusions about economic growth can then be obtained by averaging over all models and weighing each model by its posterior probability $P(m_j|y)$. This procedure is known as *Bayesian model averaging* (see Mitchell and Beauchamp (1988) and Raftery et al. (1997)) and was also used by Sala-i-Martin (1997), Fernandez et al. (2001b) and Sala-i-Martin et al. (2004) to study economic growth. Following these earlier studies, we specify prior model probabilities by assigning an independent prior inclusion probability of \bar{k}/k to each regressor, where \bar{k} is the a priori expected model size and k is the number of candidate regressors. As in Sala-i-Martin et al. (2004), we set \bar{k} equal to 7.

3.3 Posterior inference

Conditional on the median correlation $\bar{\rho}$ and the selection of variables to include in the regression, the marginal data likelihood and the parameter posterior distributions can be calculated analytically. However, considering all possible subsets of explanatory variables, the full model space now contains $2^{67} \approx 1.5 * 10^{20}$ different models, which makes it impossible to explicitly average over all candidate models. Fortunately, the work by Fernandez et al. (2001b) and Sala-i-Martin et al. (2004) suggests that the posterior probability mass is typically concentrated in a relatively small fraction of these models, making it feasible to simulate from the posterior distribution over models. To accomplish this, we use the MC³ methodology of Madigan and York (1995), which is a Metropolis-Hastings algorithm on the model space. The MC³ algorithm uses a uniform proposal distribution on the model space containing the current model and all models obtained by adding or removing a regressor. By using the stochastic Metropolis-Hastings acceptance criterion, the algorithm is ensured to have the posterior distribution over models as its stationary distribution. After each MC³ step, we draw a new proposal value for $\bar{\rho}$ from its prior in an independence chain Metropolis-Hastings step.

3.4 Data

We estimate the new nonlinear models on the data set of Sala-i-Martin et al. (2004), which is a collection of the variables examined in earlier studies. All variables in this data set were measured at the beginning of the sample period (1960) in order to avoid endogeneity problems, with the exception of the variables related to war, inflation and the openness of the economy. In addition, some countries and variables were excluded because of the need for a 'balanced' data set, i.e. a data set where all regressors are observed for all countries. After this selection process, the data set contains 88 countries, each with data on GDP growth and 67 explanatory variables. The data set contains countries in different stages of development and with a wide geographic dispersion. The explanatory variables cover a wide range of different factors, including data on economic development, social issues, health, geography, politics, education and more. The list of variables is given along with a summary of the posterior results in Section 4.2. For further discussion of the data as well as a list of sources we refer the reader to Sala-i-Martin et al. (2004).

4 Results

Eleven million draws were generated from the posterior distribution over models and over $\bar{\rho}$, of which the first million were discarded as burn-in. The remaining draws contain 120 thousand unique model specifications. By numerically integrating out $\bar{\rho}$ from the likelihood for these specifications we can compare the exact proportional posterior probabilities of these models to the number of times they were sampled (see Ley and Steel, 2009). The correlation between these measures is extremely close to one, indicating convergence of the sampling procedure. In addition, the trace plots for $\bar{\rho}$, the model size, and the different variable inclusion indicators support this conclusion.

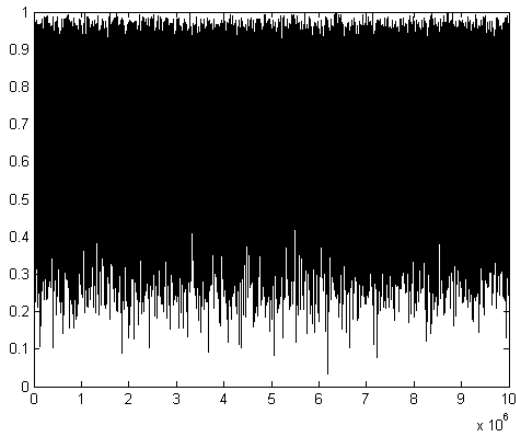


Figure 1: trace plot $\bar{\rho}$

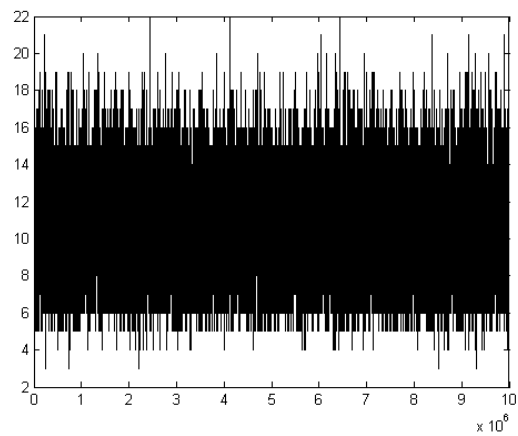


Figure 2: trace plot model size

The corresponding posterior distributions for the median cross-country correlation $\bar{\rho}$ and for the model size are displayed below.

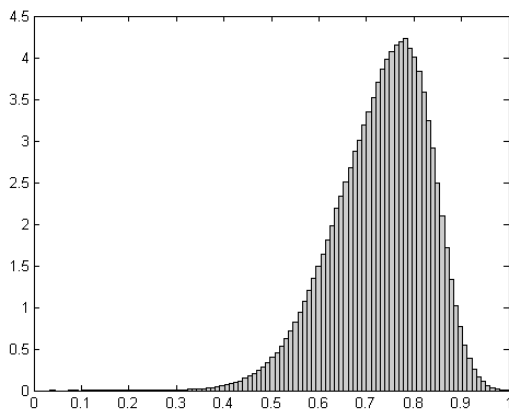


Figure 3: Posterior distribution $\bar{\rho}$

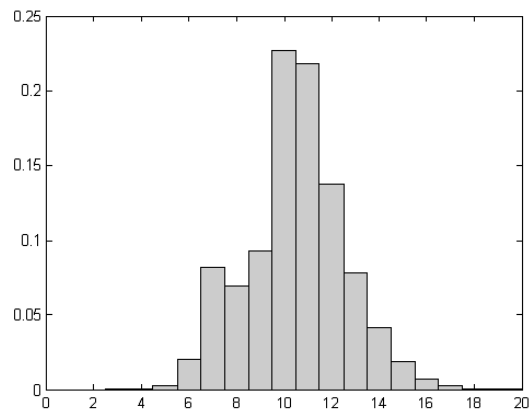


Figure 4: Posterior distribution model size

The posterior correlation between $\bar{\rho}$ and the model size is -0.027, indicating no strong relationship overall between cross-country correlation and the number of regressors in the model. However, we find that those draws with $\bar{\rho}$ very close to one do have a somewhat smaller model size. Using the methodology of George and McCulloch (1997) we find that 10 million draws is enough to cover about 70% of the posterior model probability if we keep $\bar{\rho}$ fixed. Since the correlation between $\bar{\rho}$ and the model size is low we expect to have covered a similar percentage in the analysis presented here.

4.1 Evidence for parameter heterogeneity

The posterior distribution of the median correlation $\bar{\rho}$, displayed in Figure 3, shows that the data support a median cross-country correlation of around 0.75. The posterior density is low at both $\bar{\rho} = 0$ and $\bar{\rho} = 1$, indicating that the regression coefficients are most likely not independent, but also not equal across countries. A formal test of the linear model specification against the new model can be performed by constructing a Bayes Factor $p(y|\bar{\rho} = 1)/p(y)$. By considering this Bayes Factor as a Savage-Dickey density ratio (Dickey (1971), Verdinelli and Wasserman (1995)) we can see that it reduces to $p(\bar{\rho}|y)|_{\bar{\rho}=1}$, the posterior density at $\bar{\rho} = 1$. This density can be calculated accurately using the method of Chib and Jeliazkov (2001) and it turns out to be $1.15 * 10^{-4}$, providing strong evidence in support of parameter heterogeneity. This result is consistent with the conclusions of Durlauf and Johnson (1995) and Liu and Stengos (1999) who have documented this type of nonlinearity before.

One might suspect that the nonlinear models have an unfair advantage in this comparison since they always include an (indirect) influence of the variables for initial GDP and higher education regardless of the variable selection, while the linear models do not. However, restricting the linear models to always include these variables does not substantially increase their posterior probability. In addition, we find that the posterior distribution for $\bar{\rho}$ is robust to changes in the prior parameters g and \bar{k} . For example, using $g = 1/k^2$ and $\bar{k} = k/2$ as in Fernandez et al. (2001b) produces a very similar posterior distribution for $\bar{\rho}$, with similar evidence in favor of multiple regime parameter heterogeneity. The posterior distribution of the model size is more strongly affected by changes in g and k , see Ley and Steel (2010).

4.2 Posterior summary regression coefficients

The posterior distribution of the regression coefficients β is a mixture of multivariate Student's t densities, depending on the sampled $\bar{\rho}$ and variable selections, and can be obtained analytically from equation (12). The characteristics of this posterior distribution are summarized in Table 1. The first column of the table lists the posterior inclusion probability of each variable, i.e. the sum of the posterior probabilities of all models including that variable. Conditional on the inclusion of each regressor, the table lists the expected average parameter for that regressor across countries, along with its posterior standard deviation.

This gives us a sense of the average directional effect of each explanatory variable and facilitates comparison with the earlier studies. Using the terminology of the preceding literature, a regressor can be considered robust if the bulk of the posterior mass of its average parameter lies either above or below zero. In order to determine this, the posterior confidence for the sign of the average parameter of each regressor is reported, which Sala-i-Martin et al. (2004) called the 'sign certainty probability'. Also reported are the expected number of countries with a parameter of this sign and the number of countries that have at least a 90% posterior probability of having a parameter of this sign.

The reported regressors are ordered according to posterior inclusion probability. The variables that were found to be 'significant' by Sala-i-Martin et al. (2004) are printed in bold face to facilitate comparison. These variables are the ones for which they find a higher posterior inclusion probability than the prior inclusion probability. In their results, most of these variables also have a sign certainty probability above 0.975 and are thus considered 'robust'. The variables with increased inclusion probability under our specification are separated from the others by a horizontal line.

Table 1: Parameter estimates regressors

Variable	Posterior inclusion probability	Posterior mean avg. parameter conditional on inclusion	Posterior std. dev. average parameter	Sign certainty probability	\mathbb{E} fraction of countries with parameter of this sign	Number of countries with > 90% prob of having a param. of this sign
East Asian Dummy	0.994	0.0177	0.0056	0.999	0.859	53
Political Rights	0.986	-0.0013	0.0012	0.876	0.620	23
Investment Price	0.978	-0.0001	4.22e-5	0.992	0.846	42
Life Expectancy in 1960	0.832	0.0012	0.0003	0.998	0.915	68
GDP in 1960 (log)	0.800	-0.0223	0.0049	1.000	0.902	63
Fertility in 1960s	0.777	0.0108	0.0120	0.887	0.667	31
Socialist Dummy	0.637	0.0045	0.0061	0.771	0.634	31
Malaria Prevalence in 1960s	0.541	-0.0140	0.0059	0.995	0.821	45
Gov. Consumption Share 1960s	0.423	-0.0468	0.0203	0.984	0.774	35
Fraction Speaking Foreign Language	0.386	0.0124	0.0036	0.998	0.867	58
Terms of Trade Growth in 1960s	0.273	0.0358	0.0422	0.810	0.600	22
Fraction of Tropical Area	0.222	-0.0067	0.0052	0.917	0.694	25
Real Exchange Rate Distortions	0.189	-0.0001	4.15e-5	0.995	0.809	53
Government Share of GDP in 1960s	0.172	-0.0534	0.0209	0.984	0.822	27
Primary Schooling in 1960	0.165	0.0339	0.0096	1.000	0.841	50
Spanish Colony	0.159	-0.0127	0.0063	0.978	0.819	33
Air Distance to Big Cities	0.158	-5.56e-7	1.8e-6	0.657	0.605	0
Fraction Muslim	0.154	0.0076	0.0050	0.940	0.719	37
Openness measure 1965-74	0.151	-0.0005	0.0046	0.514	0.522	14
Nominal Government GDP Share 1960s	0.142	-0.0541	0.0348	0.946	0.739	31
Public Investment Share	0.105	0.0008	0.0391	0.512	0.471	10

Table 1: Parameter estimates continued

Variable	Posterior inclusion probability	Posterior mean avg. parameter conditional on inclusion	Posterior std. dev. average parameter	Sign certainty probability	\mathbb{E} fraction of countries with parameter of this sign	Number of countries with $> 90\%$ prob of having a param. of this sign
Hydrocarbon Deposits in 1993	0.095	0.0004	0.0003	0.900	0.683	21
Latin American Dummy	0.093	-0.0062	0.0068	0.826	0.662	15
Population Density 1960	0.082	4.67e-6	1.23e-5	0.672	0.631	21
Tropical Climate Zone	0.082	-0.0006	0.0064	0.542	0.524	12
Population Density Coastal in 1960s	0.071	-8.80e-6	5.44e-6	0.912	0.645	36
Fraction Population Less than 15 Average	0.066	0.1567	0.0610	0.996	0.816	54
Inflation 1960-90	0.052	0.0001	0.0001	0.880	0.716	8
Fraction of Land Area Near Navigable Water	0.049	-0.0055	0.0067	0.774	0.634	20
Civil Liberties	0.047	-0.0174	0.0076	0.981	0.796	62
Square of Inflation 1960-90	0.044	1.81e-6	1.86e-6	0.850	0.696	4
British Colony Dummy	0.038	-0.0009	0.0034	0.613	0.554	15
African Dummy	0.036	-0.0075	0.0082	0.834	0.653	25
Fraction Hindus	0.035	0.0083	0.0172	0.705	0.635	1
Fraction Population In Tropics	0.030	-0.0142	0.0072	0.972	0.836	34
European Dummy	0.030	-0.0319	0.0113	0.998	0.760	49
Fraction Protestants	0.026	-0.0338	0.0176	0.982	0.848	54
Fraction Confucian	0.025	-0.0583	0.0378	0.937	0.714	36
Fraction GDP in Mining	0.024	-0.0045	0.0284	0.582	0.484	1
Primary Exports 1970	0.021	0.0116	0.0085	0.928	0.671	29
Absolute Latitude	0.021	-3.68e-5	0.0002	0.518	0.521	8
Ethnolinguistic Fractionalization	0.019	-0.0060	0.0062	0.835	0.653	5
Oil Producing Country Dummy	0.019	-0.0097	0.0082	0.882	0.684	1
Years Open 1950-94	0.019	0.0038	0.0070	0.703	0.551	0
Fraction Othodox	0.018	-0.0336	0.0266	0.897	0.697	21
Fraction Catholic	0.017	-0.0045	0.0071	0.750	0.581	18
Population in 1960	0.016	-4.66e-8	3.66e-8	0.822	0.650	0
Fraction Spent in War 1960-90	0.016	0.0069	0.0104	0.751	0.621	10
Fraction Buddhist	0.016	0.0020	0.0119	0.558	0.549	0
Population Growth Rate 1960-90	0.015	0.0721	0.5810	0.568	0.527	2
Fraction Population Over 65	0.015	-0.1503	0.1430	0.878	0.654	20
Timing of Independence	0.012	0.0006	0.0020	0.592	0.570	13
Landlocked Country Dummy	0.012	-0.0119	0.0040	0.996	0.712	39
Higher Education 1960	0.010	-0.2731	0.0913	1.000	0.854	59
Defense Spending Share	0.009	0.0178	0.0696	0.574	0.516	3
War Participation 1960-90	0.008	0.0018	0.0028	0.733	0.622	0
Land Area	0.008	-2.81e-9	1.30e-9	0.827	0.637	0
Size of Economy	0.008	-0.0005	0.0018	0.545	0.552	0
Capitalism	0.008	-0.0005	0.0012	0.670	0.557	4
Outward Orientation	0.006	-0.0034	0.0025	0.914	0.700	24
English Speaking Population	0.006	-0.0220	0.0168	0.930	0.760	39
Colony Dummy	0.005	0.0053	0.0044	0.890	0.681	22
Revolutions and Coups	0.004	-0.0011	0.0060	0.581	0.536	0
Interior Density	0.004	-5.68e-6	2.28e-5	0.565	0.508	17

Table 1: Parameter estimates continued

Variable	Posterior inclusion probability	Posterior mean avg. parameter conditional on inclusion	Posterior std. dev. average parameter	Sign certainty probability	\mathbb{E} fraction of countries with parameter of this sign	Number of countries with > 90% prob of having a param. of this sign
Public Education Spending Share in GDP in 1960s	0.004	0.1203	0.1678	0.761	0.617	20
Terms of Trade Ranking	0.003	7.41e-5	0.0125	0.524	0.510	0
Religion Measure	0.003	-0.0010	0.0098	0.548	0.503	0

The results of our model averaging confirm the general conclusion of Fernandez et al. (2001b) and Sala-i-Martin et al. (2004), indicating the importance of a number of the same variables. Most of these variables also have average parameter estimates of the same sign as those in the earlier studies. However, we also find some important differences. The most striking difference is that in the earlier studies there was a very strong positive correlation between inclusion probability and sign certainty. For example, in the analysis of Sala-i-Martin et al. (2004) the twenty variables with the highest inclusion probability are also the twenty variables with the highest sign certainty. In our analysis this is very different: a number of variables have a high inclusion probability despite having low sign certainty and several variables with high sign certainty have a low posterior inclusion probability. On average, the sign certainty is slightly decreased. This difference is a direct result of allowing the parameters of the variables to differ over countries. In the linear model specification only those variables that have a similar effect across countries are likely to be included, while in our specification also variables with heterogeneous effects can have predictive power.

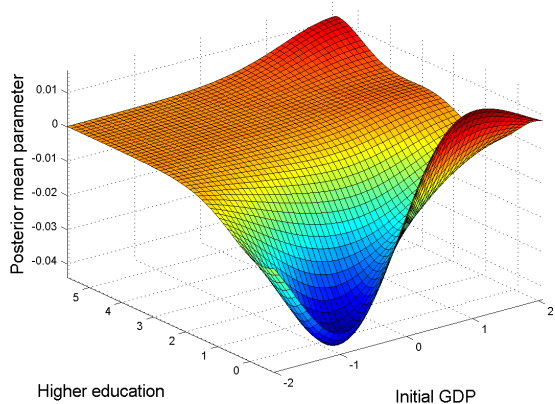
In addition to a somewhat reduced sign certainty for the average parameters, the two last columns of the table show that the parameter estimates for each country individually are much less certain than under the linear model specification and that the expected signs of the parameters may differ strongly over countries. This means that many of the variables found to be robust in earlier studies do not have robust marginal effects across countries and variable selections under our model specification. Although this is to be expected when allowing the variables to be country-dependent, it contradicts the findings of Minier (2007) who finds that allowing for nonlinearity makes the parameter estimates more robust. However, as already discussed in Section 2, those results are based on only the linear components of the model and not on the full marginal relationships which makes comparison difficult. Interestingly, we do confirm the finding of Minier (2007) that variables related to fiscal policy become more important (in terms of inclusion probability) when allowing for nonlinearity. Examples of this are the *government consumption share*, the *government share of GDP*, and the *public investment share*.

As in the earlier studies, the variable with the highest posterior inclusion probability is the dummy for *East Asian countries*. The influence of this dummy variable reflects the high rate of economic growth in this region during the sample period. This dummy variable has a high correlation with the *fraction of the population that is Confucian*, which also belongs to the significant variables of Sala-i-Martin et al. (2004). However, with the current

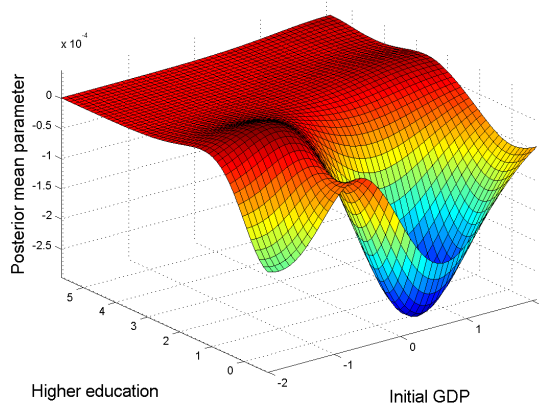
specification the East Asian dummy seems to be sufficient to explain the growth of the countries in this region, giving the Confucian variable a much lower inclusion probability. The East Asian dummy is a good example of the importance of geographic location to economic growth, as are the variables for the *fraction of tropical area* and the *air distance to big cities*. Also consistent with the earlier studies are the findings that education (*primary schooling*), investment (*investment price, public investment share*) and health (*life expectancy, Malaria prevalence*) are important to economic growth.

4.3 Parameter posterior means

Since the parameters in our specification are allowed to vary with GDP and higher education, the posterior means of the parameters are functions of these initial conditions. These posterior mean functions are summarized in Table 4.2, but this summary does not fully describe the heterogeneity present in the posterior mean functions of the parameters. To offer a more detailed look at this heterogeneity, this section shows and discusses the full posterior mean functions of selected regressors, conditional on inclusion. The variance around these posterior mean functions is reasonably high in most cases, so the comments below are mostly illustrative.



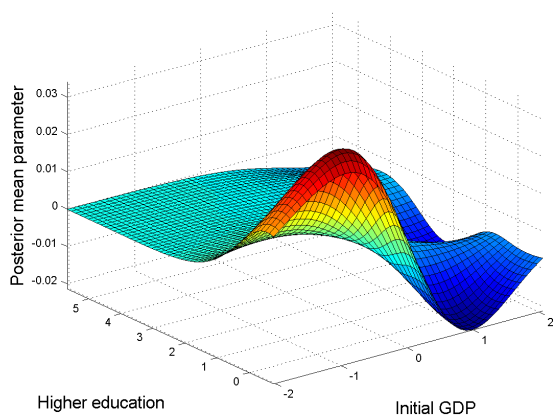
Initial GDP



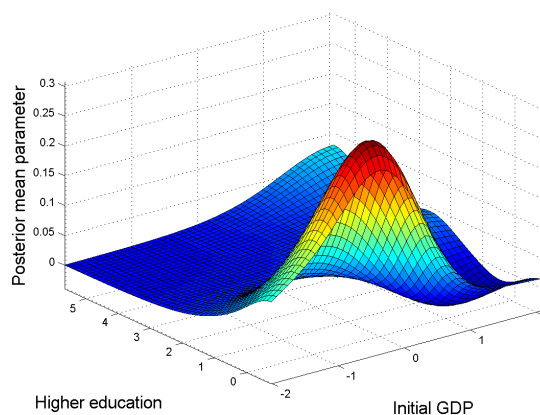
Investment Price

The first posterior mean function shown is that for the *initial log GDP*. The graph shows that the marginal effect of initial GDP on growth is largely negative, but that as initial GDP rises the marginal effect rises with it. This shows that the nonlinearity in this variable, found in several earlier studies, does not come only from its interaction with the other variables but also through its direct relationship with economic growth. In addition, the marginal effect of initial GDP seems to go up with the level of education, indicating that rich countries did better when they had a high level of education.

The second parameter shown concerns the *investment price*. The price of investment is important in determining economic growth, judging by the high posterior inclusion probability for this variable. High costs of investment are detrimental to economic growth as indicated by the negative posterior mean for the coefficient of this variable and its high sign certainty. The posterior mean graph shows that this negative effect is present for most of the countries in the sample.

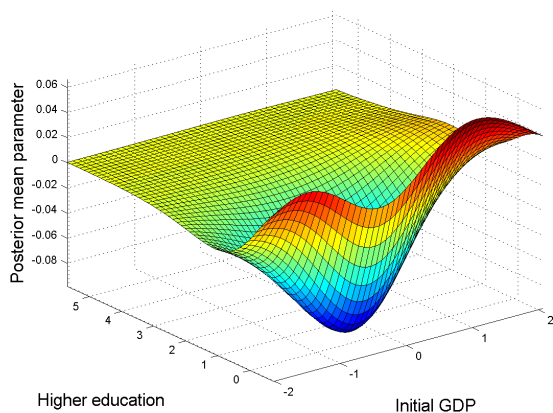


Fertility

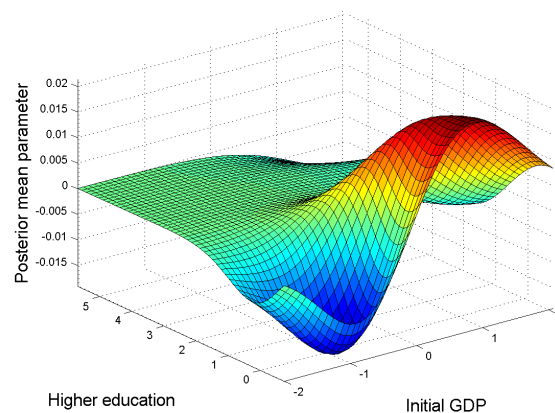


Fraction population under 15

The posterior means of the parameters for the *fertility* variable and the *fraction of the population under 15* have similar patterns, indicating that these variable might represent the same underlying effect. The positive posterior means for the average parameters for these variables suggest that having a young population promotes economic growth, presumably because it means that these countries had a large labor force during the sample period. This is further supported by the negative posterior mean for the average coefficient of the *fraction of the population over 65*.



Fraction GDP in mining



Openness economy

An interesting difference between our results and those of Sala-i-Martin et al. (2004) is the sign of the posterior mean of the average parameter for *fraction GDP in mining*. They find a positive relationship with high sign certainty while we estimate the average parameter to be negative. The latter is consistent with the popular idea that large endowments of natural resources are associated with rent-seeking and low growth, known as the 'natural resources curse' (see Sachs and Warner, 2001). This difference can be explained by the strong parameter heterogeneity shown in the graph above. While most countries with large endowments of natural resources indeed seem to grow more slowly, a few countries with high mining activity actually experienced very high growth rates. Examples include Chile and Saudi Arabia, which both had high initial GDP and low education, explaining the positive posterior mean in the right corner of the graph. The posterior mean is also very high for countries with low education and low initial GDP. This effect is most likely caused in large part by Botswana, the country with the highest share of GDP from mining in the sample. This country has successfully managed to exploit the diamonds found in its territory and has experienced extremely high growth rates. This result provides a powerful argument in favor of parameter heterogeneity as this finding is clearly not just a statistical anomaly: Botswana, Chile and Saudi Arabia really do benefit from their natural resources, while for other countries the case for the natural resources curse is quite strong.

The *openness of the economy* also has a very interesting pattern of parameter heterogeneity. The graph of the posterior mean suggests that having an open economy was especially beneficial for those countries that were already comparatively rich. Poor countries seem to have derived much less benefit from having an open economy as their parameter posterior means are negative.

5 Conclusion

A rigorous regression analysis of cross-country growth data should jointly take into account the model uncertainty present in the variable selection problem as well in the functional form specification. Although both sources of model uncertainty have separately received much attention in the literature, joint treatments are unfortunately very rare. This paper presents such an integrated analysis.

We address the model uncertainty relating to the functional form of the regression function by introducing a new flexible growth regression model based on a Gaussian process prior. The new model explicitly allows for multiple regime parameter heterogeneity as suggested by new growth theory, while nesting the linear model specification as a special case. We solve the variable selection problem by performing a Bayesian model averaging with the new model using different sets of explanatory variables. As argued earlier by Fernandez et al. (2001b) and Sala-i-Martin et al. (2004), this approach provides a theoretically sound and practical way of considering a large class of different variable selections.

A formal model comparison provides strong evidence supporting multiple regime parameter heterogeneity, consistent with the conclusions of Durlauf and Johnson (1995) and Liu and Stengos (1999) who have documented this type of nonlinearity before. The results do not support the conclusions of Minier (2007) who finds that allowing for nonlinearity makes the parameter estimates more robust to the variable selection. In addition, our results show that many of the explanatory variables do not have robust partial correlations to growth across countries.

The proposed method is general and can easily be adapted to allow for almost any kind of nonlinearity in the functional form of the regression function. Exploring these different specifications is a promising direction for future research.

References

- Aghion, P., E. Caroli, and C. Garcia-Penalosa (1999). Inequality and economic growth: The perspective of the new growth theories. *Journal of Economic Literature* 37(4), 1615–1660.
- Barro, R. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics* 106(2), 407–43.
- Basturk, N., R. Paap, and D. van Dijk (2010). Structural differences in economic growth: An endogenous clustering approach. *Applied Economics* forthcoming.
- Chib, S. and I. Jeliazkov (2001). Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association* 96, 270–281.
- Cuaresma, C. and G. Doppelhofer (2007). Nonlinearities in cross-country growth regressions: A Bayesian averaging of thresholds (BAT) approach. *Journal of Macroeconomics* 29, 541–554.
- Dickey, J. M. (1971). The weighted likelihood ratio, linear hypotheses on normal location parameters. *The Annals of Mathematical Statistics* 42(1), 204–223.
- Durlauf, S. N. and P. A. Johnson (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10(4), 365–384.
- Fernandez, C., E. Ley, and M. F. Steel (2001a). Benchmark priors for Bayesian model averaging. *Journal of Econometrics* 100(2), 381–427.
- Fernandez, C., E. Ley, and M. F. Steel (2001b). Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics* 16(5), 563–576.
- George, E. I. and R. E. McCulloch (1997). Approaches for Bayesian variable selection. *Statistica Sinica* 7, 339–373.
- Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics*. Wiley-Interscience.
- Grier, K. B. and G. Tullock (1989). An empirical analysis of cross-national economic growth. *Journal of Monetary Economics* 24(2), 259–276.
- Kalaitzidakis, P., T. P. Mamuneas, and T. Stengos (2000). A non-linear sensitivity analysis of cross-country growth regressions. *Canadian Journal of Economics* 33(3), 604–17.
- Koop, G. and D. J. Poirier (2004). Bayesian variants of some classical semiparametric regression techniques. *Journal of Econometrics* 123(2), 259–282.
- Kormendi, R. and P. Meguire (1985). Macroeconomic determinant of growth, cross-country evidence. *Journal of Monetary Economics* 16, 141–163.

- Leamer, E. E. (1983). Let's take the con out of econometrics. *American Economic Review* 73(1), 31–43.
- Levine, R. and D. Renelt (1992). Sensitivity analysis of cross-country growth regressions. *American Economic Review* 82(4), 942–963.
- Ley, E. and M. F. Steel (2009). On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics* 24, 651–674.
- Ley, E. and M. F. Steel (2010). Mixtures of g-priors for Bayesian model averaging with economic applications.
- Liu, Z. and T. Stengos (1999). Non-linearities in cross country growth regressions: A semiparametric approach. *Journal of Applied Econometrics* 14(5), 527–538.
- Maasoumi, E., J. Racine, and T. Stengos (2007). Growth and convergence: A profile of distribution dynamics and mobility. *Journal of Econometrics* 136(2), 483 – 508.
- Madigan, D. and J. York (1995). Bayesian graphical models for discrete data. *International Statistical Review* 63, 215–232.
- Mankiw, N., D. Romer, and D. Weil (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107(2), 407–437.
- Minier, J. (2007). Nonlinearities and robustness in growth regressions. *American Economic Review* 97(2), 388–392.
- Mitchell, T. and J. Beauchamp (1988). Bayesian variable selection in linear regression. *Journal of the American Statistical Association* 83, 1023–1046.
- Paap, R., P. Franses, and D. van Dijk (2005). Does Africa grow slower than Asia, Latin America and the Middle East? evidence from a new data-based classification method. *Journal of Development Economics* 77, 553–570.
- Raftery, A., D. Madigan, and J. Hoeting (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association* 92, 179–191.
- Rasmussen, C. E. and C. K. I. Williams (2006). *Gaussian Processes for Machine Learning*. MIT Press.
- Sachs, J. D. and A. M. Warner (2001). The curse of natural resources. *European Economic Review* 45(4-6), 827–838.
- Sala-i-Martin, X. (1997). I just ran two million regressions. *American Economic Review* 87(2), 178–183.

- Sala-i-Martin, X., G. Doppelhofer, and R. I. Miller (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *American Economic Review* 94(4), 813–835.
- Verdinelli, I. and L. Wasserman (1995). Computing Bayes Factors using a generalization of the Savage-Dickey density ratio. *Journal of the American Statistical Association* 90, 614–618.
- Zellner, A. (1986). On assessing prior distributions and Bayesian regression analysis with g -prior distributions. In P. Goel and A. Zellner (Eds.), *Bayesian Inference and Decision Making: Essays in Honour of Bruno de Finetti*, pp. 233–243. North-Holland, Amsterdam.