Available online at www.sciencedirect.com
ScienceDirect
Journal of Macroeconomics 29 (2007) 476–493
Journal of MACROECONOMICS
www.elsevier.com/locate/jmacro

Jointness in Bayesian variable selection with applications to growth regression

Eduardo Ley [a,*], Mark F.J. Steel [b,1]

[a] *The World Bank, 1818 H Street NW, Washington DC 20433, USA*
[b] *Department of Statistics, University of Warwick, Coventry CV4 7AL, UK*

Received 6 October 2006; accepted 19 December 2006
Available online 8 April 2007

Abstract

We present a measure of jointness to explore dependence among regressors, in the context of Bayesian model selection. The jointness measure proposed here equals the posterior odds ratio between those models that include a set of variables and the models that only include proper subsets. We illustrate its application in cross-country growth regressions using two datasets from the model-averaging growth literature.

JEL classification: C7; C11

Keywords: Bayesian model averaging; Complements; Model uncertainty; Posterior odds; Substitutes

1. Introduction

Performing inference on the determinants of GDP growth is challenging because, in addition to the complexity and heterogeneity of the objects of study, a key characteristic of the empirics of growth lies in its open-endedness (Brock and Durlauf, 2001). Open-endedness entails that, at a conceptual level, alternative theories may suggest additional determinants of growth without necessarily excluding determinants proposed by other theories.

* Corresponding author. Tel.: +1 202 458 5089.
E-mail addresses: eley@worldbank.org (E. Ley), M.F.Steel@stats.warwick.ac.uk (M.F.J. Steel).
[1] Tel.: +44 (0) 24 7652 3369.

0164-0704/$ - see front matter
doi:10.1016/j.jmacro.2006.12.002

The absence, at the theoretical level, of such tradeoff leads to substantial model uncertainty, at the empirical level, about which variables should be included in a growth regression. In practice, a substantial number of growth determinants may be included as explanatory variables. If two such variables are capturing different sources of relevant information and should both be included, we will talk of *jointness* (as defined later), whereas if they perform very similar roles they should not appear jointly, which we will denote by *disjointness*. We could think of these situations as characterized by the covariates being complements or substitutes, respectively.

Various approaches to deal with this model uncertainty have appeared in the literature: early contributions are the extreme-bounds analysis in Levine and Renelt (1992) and the confidence-based analysis in Sala-i-Martin (1997). Fernández et al. (2001b, FLS henceforth) use Bayesian model averaging (BMA, see Hoeting et al., 1999) to handle the model uncertainty that is inherent in growth regressions, as discussed above. BMA naturally deals with model uncertainty by averaging posterior inference on quantities of interest over models, with the posterior model probabilities as weights. Other papers using BMA in this context are León-González and Montolio (2004) and Papageorgiou and Masanjala (2005). Alternative ways of dealing with model uncertainty are proposed in Sala-i-Martin et al. (2004, SDM henceforth),[2] and Tsangarides (2005). Insightful discussions of model uncertainty in growth regressions can be found in Brock and Durlauf (2001) and Brock et al. (2003). All of these studies adopt a Normal linear regression model and consider modeling n growth observations in y using an intercept and explanatory variables from a set of k variables in Z, allowing for any subset of the variables in Z to appear in the model. This results in 2^k possible models, which will thus be characterized by the selection of regressors. We call model M_j the model with the $0 \leqslant k_j \leqslant k$ regressors grouped in Z_j, leading to

$$y|\alpha, \beta_j, \sigma \sim N(\alpha \iota_n + Z_j \beta_j, \sigma^2 I), \tag{1}$$

where ι_n is a vector of n ones, $\beta_j \in \Re^{k_j}$ groups the relevant regression coefficients and $\sigma \in \Re_+$ is a scale parameter.

Based on theoretical considerations and simulation results in Fernández et al. (2001a), FLS adopt the following prior distribution for the parameters in M_j:

$$p(\alpha, \beta_j, \sigma|M_j) \propto \sigma^{-1} f_N^{k_j}(\beta|0, \sigma^2 (g Z_j' Z_j)^{-1}), \tag{2}$$

where $f_N^q(w|m, V)$ denotes the density function of a q-dimensional Normal distribution on w with mean m and covariance matrix V and they choose $g = 1/\max\{n, k^2\}$. Finally, the components of β not appearing in M_j are exactly zero, represented by a prior point mass at zero.

The prior model probabilities are specified by $P(M_j) = \theta^{k_j}(1-\theta)^{k-k_j}$, which implies that each regressor enters a model independently of the others with prior probability θ. Thus, the prior expected model size is θk. We follow Fernández et al. (2001a) and FLS in choosing $\theta = 0.5$, which is a benchmark choice—implying that $P(M_j) = 2^{-k}$ and that expected model size is $k/2$. Throughout this paper, we shall use the same prior as in FLS. An explicit analysis of alternative priors in this context is carried out in Ley and Steel (2007).

[2] SDM's procedure BACE in fact uses approximate Bayesian posterior probabilities of regression models based on the Schwarz criterion, as proposed by Raftery (1995).