Determining Growth Determinants: Default Priors and Predictive Performance in Bayesian Model Averaging *

Theo S. Eicher
Department of Economics
University of Washington

Chris Papageorgiou
Research Department
International Monetary Fund

Adrian E. Raftery
Departments of Statistics and Sociology
Center for Statistics and the Social Sciences
University of Washington

August, 2007

Abstract

Economic growth has been a showcase of model uncertainty, given the many competing theories and candidate regressors that have been proposed to explain growth. Bayesian Model Averaging (BMA) addresses model uncertainty as part of the empirical strategy, but its implementation is subject to the choice of priors: the priors for the parameters in each model, and the prior over the model space. For a well-known growth dataset, we show that model choice can be sensitive to the prior specification, but that economic significance (model-averaged inference about regression coefficients) is quite robust to the choice of prior. We provide a procedure to assess priors in terms of their predictive performance. The Unit Information Prior, combined with a uniform model prior outperformed other popular priors in the growth dataset and in simulated data. It also identified the richest set of growth determinants, supporting several new growth theories. We also show that there is a tradeoff between model and parameter priors, so that the results of reducing prior expected model size and increasing prior parameter variance are similar. Our branch-and-bound algorithm for implementing BMA was faster than the alternative coin flip importance sampling and MC3 algorithms, and was also more successful in identifying the best model.

JEL Classification: O51, O52, O53.

Keywords: Growth Determinants; Model Uncertainty; Bayesian Model Averaging (BMA); Parameter and Model Prior Elicitation; Predictive Performance.

*We are grateful to Amanda Cox for her tireless support, advice, and programming. We thank Drew Creal for excellent software programming, Tilmann Gneiting for kindly sharing his CPRS code for BMA applications, and Eduardo Ley for sharing data. We also thank Veronica Berrocal, Gernot Doppelhofer, Edward George, Tilmann Gneiting, Jennifer Hoeting, Andros Kourtellos, Andreas Leukert, Eduardo Ley, Chih Ming Tan, and seminar participants at the Department of Statistics, University of Washington for helpful comments and discussions. Fred Nick at the University of Washington Center for Social Science Computation and Research provided crucial computing support. Eicher gratefully acknowledges financial support from the University of Washington Center for Statistics and the Social Sciences through a seed grant. Raftery's research was supported by the DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research under Grant N00014-01-10745. Raftery thanks Miroslav Kárny and the Department of Adaptive Systems, Institute for Information Theory and Automation, Prague, Czech Republic, as well as Gilles Celeux and INRIA Futurs, France, for hospitality during the preparation of this paper. The views expressed in this study are the sole responsibility of the authors and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

1 Introduction

Economic growth has been a showcase of model uncertainty since the early 1990s when a surge of new growth theories gave rise to a large literature that sought to evaluate the new growth determinants (see Durlauf, Johnson and Temple, 2005 for a survey). Recent growth literature has used Bayesian Model Averaging (BMA), which is specifically designed to address model uncertainty as part of the empirical strategy (e.g., Fernández, Ley and Steel, 2001a; Sala-i-Martin, Doppelhofer and Miller, 2004). The advantages of BMA are that it incorporates model uncertainty and also assesses the robustness of conclusions to model assumptions in a principled way. The implementation of BMA is, however, subject to a major challenge (and criticism): it requires prior distributions over all parameters in all models, and the prior probability of each model must also be specified. In this paper we examine the sensitivity of growth determinants to alternative prior specifications, develop a procedure to evaluate competing priors, and suggest a default prior that provided consistently good results in our experiments.

Previous applications of BMA to economic growth highlighted its ability to account for model uncertainty but did not emphasize that prior distributions may well influence results. With well defined parameters and large sample sizes, reasonable choices of prior distributions have only minor effects on posterior inferences (Leamer, 1978). Unfortunately, the definitions of 'well identified' and 'large' sample size are often problematic in economics. Datasets in economics, and especially in growth applications, are generally small (often less than 100 observations) and the number of candidate regressors motivated by theory can be large (with over 40 regressors). The large number of models and the small number of observations pose challenges for both prior specification and computation. In statistics, it is thus a common practice to assess the potential sensitivity of inferences to changes in the prior distributions. In economics, such systematic sensitivity analysis is not the norm, nor is it part of the empirical toolbox.[1]

We provide an integrated evaluation procedure to examine the impact of subjective and objective prior distributions on BMA inference. The procedure allows for the direct comparison of a dozen popular parameter priors in addition to any given prior over the model space. Since different prior structures may imply different relationships between regressors and dependent variables, our procedure also evaluates results, based on predictive performance. To date, such an integrated evaluation procedure that simultaneously compares model and parameter prior along with their

[1] Recent examples of prior robustness analyses in economic applications include Sala-i-Martin, Doppelhofer and Miller (2004), Durlauf, Kourtellos and Tan (2006; 2007), and Ley and Steel (2007b).

predictive performance has not been offered.

Using a prominent growth dataset that contains 41 regressors compiled by Fernández, Ley and Steel (2001a) (hereafter FLS), we show that the number of growth determinants implied by the various popular priors varies widely – from as few as 3 to as many as 22. We suggest it would be ill-advised to search for common regressors among alterative priors and declare them robustly related to growth. Instead, our evaluation criterion is out-of-sample predictive performance, which provides a neutral basis for comparing methods. We assess predictive performance in terms of point estimation (the Mean Squared Error), and in terms of the quality of the predictive probability distribution as a whole. Our measures of the latter take account of both sharpness (narrow prediction intervals) and calibration (probability estimates that are correct). We introduce the Continuous Ranked Probability Score (CRPS) to economics, which captures both sharpness and calibration (Matheson and Winkler, 1976), but which is less sensitive to outliers and extreme events than alternative measures that have been used in the previous economics literature, such as the Log Predictive Score (Weigend and Shi, 2000; Gneiting and Raftery, 2007).

In our experiments, we found that the Unit Information Prior (UIP) for the parameters of each model, combined with the uniform model prior, provided consistently superior out-of-sample predictive performance. This was true for our growth dataset and for simulated data. If a comparison of all possible priors is too time consuming, we therefore suggest UIP together with uniform model priors as a default prior for BMA in linear regression. From a practical point of view this is a convenient finding, because the UIP leads to a very simple approximation to the posterior model probabilities in terms of the Bayesian Information Criterion (BIC; Schwarz, 1978; Raftery, 1995) which reduces BMA's computational intensity.

The UIP performed better than the "automatic" parameter prior suggested by FLS (2001b) in the growth application and in simulated datasets. In terms of model priors, we found that the UIP, together with a uniform prior on the model space, also produced better predictive performance than the subjective model prior suggested by Sala-i-Martin, Doppelhofer and Miller (2004) (hereafter SDM). The UIP default prior not only generated consistently better predictive performance, but it also identified a larger number of growth determinants than previous analyses. Aside from Confucius (the eastern religion dummy), Initial GDP, Life Expectancy, Rule of Law and Equipment Investment, we also found robust effects of Primary and Secondary Education, Ethnolinguistic Fragmentation, Civil Liberties, Black Market Premium and Outward Orientation on growth. There has been doubt about whether the growth dataset features a sufficiently large number of observations to provide a rich set of growth determinants. Our analysis shows that, with the appropriate prior, a rich set of growth determinants provides good predictive performance.

BMA poses the challenge of computing the model average, since the number of candidate models is often huge. In our experiments we compared three popular computational methods: the *branch-and-bound* method used in our algorithm (Raftery, 1995), the *MC3* (*Markov Chain Monte Carlo Model Composition)* of Madigan and York (1995) used by FLS, and the *coin flip importance sampling* method suggested by SDM. We found that the branch-and-bound method was faster and more likely to find the model with the highest posterior probability than the other two methods.

The economics literature has long recognized model uncertainty as a central problem in regression analyses in general and in growth applications in particular. The initial approach to model selection was to use stepwise regression (Efroymson, 1960). Leamer (1978) suggested extreme bounds analysis to account not only for within-model uncertainty, but also for between-model uncertainty, which is associated with model selection.[2] The BMA methodology was developed by Leamer (1978), Raftery (1988) who coined the name, Raftery (1993), George and McCulloch (1993), Madigan and Raftery (1994) and others; for a survey of its early development see Hoeting, Madigan, Raftery and Volinsky (1999).[3] Early applications of BMA in economics include FLS (2001a), Brock and Durlauf (2001) and SDM. FLS (2001a) applied a "benchmark prior" (FLS 2001b) to the growth context, but did not report robustness analysis. Brock and Durlauf (2001) applied BMA to highlight parameter heterogeneity in popular growth datasets. SDM (whose working paper version dates back to 2000) highlighted the importance of the model prior distributions in BMA growth analysis.[4]

The paper is organized as follows. Section 2 discusses the basics of BMA theory and estimation with particular emphasis on prior structures. Section 3 presents the growth results using our integrated evaluation approach that provides an assessment of prior structures via predictive

---

[2]See Levine and Renelt (1992) and Sala-i-Martin (1997) for applications of extreme bound analysis to growth.

[3]The combination of estimates and forecasts from different models had earlier been discussed, for example by Bates and Granger (1969), Newbold and Granger (1974), Moulton (1991) and Palm and Zellner (1992), but this was in the context of equal weighting or inverse variance weighting, not of Bayesian model averaging, and was for point estimates only, not distributions.

[4]Subsequent examples of the rapidly growing literature on economics applications of BMA include policy evaluations (e.g. Brock, Durlauf, and West, 2003; and Sirimaneetham and Temple, 2006) monetary policy (e.g. Levin and Williams, 2003), macroeconomic forecasting (e.g. Garratt, Lee, Pesaran and Shin, 2003), international trade (e.g. Eicher, Henn, and Papageorgiou, 2007), environmental economics (e.g., Begun and Eicher, 2006), output volatility (e.g. Malik and Temple, 2006), and economic growth (e.g., Min and Zellner, 1993; Leon-Gonzalez and Montolio, 2004; Durlauf, Kourtellos and Tan, 2006; 2007; Eicher, Papageorgiou and Roehn, 2007; Masanjala and Papageorgiou, 2007a,b; and Ley and Steel, 2007a,b). The BMA economic forecasting literature is surveyed by Stock and Watson (2006).

performance. Section 4 confirms our results using simulated data, and Section 5 concludes.

2 Bayesian Model Averaging

In this section we briefly sketch the basic ideas of BMA and the challenges involved in implementing it, namely the choice of prior distributions for the parameters and the models, and computation of the model average. For more complete surveys of Bayesian model averaging see Raftery, Madigan and Hoeting (1997), Hoeting et al. (1999), Clyde and George (2004) and Doppelhofer (2007).

2.1 Basic BMA Methodology

BMA is a standard Bayesian solution to model uncertainty, and consists of basing prediction and inference on a weighted average over all the models considered, rather than on one single regression model. BMA requires a prior probability of each model and a prior probability distribution over the parameters of each model. The model and prior probabilities are then used to derive weights to average over all models. The approach has the attractive feature that it directly addresses questions that are central to the researcher's interests, such as "what is the probability that a model is correct?" (given that one of the models considered is), and "how likely is it that a regressor has an effect on the dependent variable?"[5]

For linear regression models, the basic setup is as follows. Given a dependent variable, Y, a number of observations, n, and a set of candidate regressors, $X_1, \dots, X_p$, the variable selection problem is to find the "best" model, or subset of regressors. We denote by $M_1, \dots, M_K$ the models considered, where each one represents a subset of the candidate regressors. Often all possible subsets are considered, in which case $K = 2^p$. Model M_k has the form

$$Y = \alpha + \sum_{j=1}^{p_k} \beta_j^{(k)} X_j^{(k)} + \varepsilon, \tag{1}$$

where $X_1^{(k)}, \dots, X_{p_k}^{(k)}$ is a subset of $X_1, \dots, X_p$, $\beta^{(k)} = (\beta_1^{(k)}, \dots, \beta_{p_k}^{(k)})$ is a vector of regression coefficients to be estimated, and $\varepsilon \sim N(0, \sigma^2)$ is the error term. We denote by $\theta_k = (\alpha, \beta^{(k)}, \sigma)$ the vector of parameters in M_k.

[5]Here we use the phrase "has an effect on the dependent variable" as shorthand for "is associated with the dependent variable after controlling for the other regressors." Even if a regressor has an effect on the dependent variable in this sense, a causal relationship is not established because the statistical result could have other sources, such as selection bias or omitted regressors. In this paper we ignore these issues, as has been common in the growth literature, and focus on the variable selection issues.

The likelihood function of model M_k, $pr(D|\theta_k,\, M_k)$, summarizes all the information about θ_k that is provided by the data, D. The *integrated likelihood* is the probability of the data given model M_k, equal to the likelihood times the prior density, $pr(\theta_k|M_k)$, integrated over the parameter space, so that

$$pr(D|M_k) = \int pr(D|\theta_k,\, M_k) pr(\theta_k|M_k) d\theta_k. \tag{2}$$

Equation (1) follows from the law of total probability.

The integrated likelihood is the crucial ingredient in deriving the model weight for model averaging. We denote by $pr(M_k)$ the prior probability that M_k is the correct model, given that one of the models considered is. Then, by Bayes's theorem, the *posterior model probability* of M_k, $pr(M_k|D)$, is equal to the model's share in the total posterior mass,

$$pr(M_k|D) = \frac{pr(D|M_k)pr(M_k)}{\sum_{\ell=1}^{K} pr(D|M_\ell)pr(M_\ell)}. \tag{3}$$

The posterior mean and variance of a regression coefficient, β_j, are then given by

$$E[\beta_j|D] = \sum_{k=1}^{K} \hat{\beta}_j^{(k)} pr(M_k|D), \tag{4}$$

$$Var[\beta_j|D] = \sum_{k=1}^{K} \left(Var[\beta_j|D,\, M_k] + (\hat{\beta}_j^{(k)})^2 \right) pr(M_k|D) - E[\beta_j|D]^2, \tag{5}$$

where $\hat{\beta}_j^{(k)}$ is the posterior mean of β_j under model M_k, and is equal to zero if X_j is not included in M_k (Raftery, 1993).

Hence the posterior mean is the weighted average of the model-specific posterior means, where the weights are equal to the models' posterior probabilities. The posterior variance reflects both the weighted average of the within-model posterior variances, and the between-model variation of the model-specific posterior means. Conditioning on a single model leaves out the between-model variation, and thus underestimates overall uncertainty. In a decision-making context, this would lead to decisions that are riskier than the decision-maker thinks they are. BMA incorporates model uncertainty into the posterior distribution itself, and thus allows it to be propagated through to final conclusions and decisions.

In addition to the posterior means and standard deviations, BMA provides the *posterior inclusion probability* of a candidate regressor, $pr(\beta_j \neq 0|D)$, by summing the posterior model probabilities across those models that include the regressor. Posterior inclusion probabilities provide a probability statement regarding the importance of a regressor that directly addresses what is often the researcher's prime concern: "what is the probability that the regressor has an effect on the dependent variable?"

BMA involves averaging over all the models considered. This can be a very large number; for example, the growth dataset we consider below features 41 candidate regressors (and so $K = 2^{41}$, or about two trillion models). Such a vast model space involves a major computational challenge; the obvious method, direct evaluation, is typically not feasible. Three practical approaches have been advocated and we comment on their efficiency below. The first is a method developed by Raftery (1995), based on the branch-and-bound algorithm of Furnival and Wilson (1974), that is guaranteed to find the single best model contained in the data.[6] The second approach is the Markov Chain Monte Carlo Model Composition (MC3) algorithm (Madigan and York, 1995), applied to BMA for regression by Raftery, Madigan and Hoeting (1997) and FLS (2001b). While efficient, MC3 does not guarantee finding the global maximum. The third approach is the coin flip importance sampling method used by SDM, whose computational efficiency is not much lower than that of MC3 if the sampling probability equals the prior inclusion probability (Clyde, DeSimone and Parmigiani, 1996). The sampling probability does not equal the prior inclusion probability in SDM's "stratified" coin flip importance sampling method and their algorithm guarantees neither finding the global maximum, nor finding the same best model in different runs.

2.2 Prior Distributions of Parameters

The implementation of BMA in linear regression as described by equations (1)-(3) is subject to a major challenge (and criticism): prior distributions must be specified over all parameters in all models. Prior probabilities of all models must also be specified. If the researcher has information about the parameters, ideally this should be reflected in the priors, and informative priors should be used, as was done, for example, by Jackman and Western (1994).

However, often the amount of prior information is small and the effort needed to specify it in terms of a probability distribution is large. Thus there have been many efforts to specify default priors that could reasonably be used for all such analyses. These are sometimes called "nonin-

[6] The algorithm organizes the model space in a tree-like way and chops off branches of models with low posterior probability. It is fast for less than about 45 regressors and for more than that tends to be slower. It is implemented for the Unit Information Prior in the BICREG function that is part of the BMA R package, available on CRAN at http://cran.r-project.org (Raftery, 1995; Raftery, Painter and Volinsky, 2005). Yeung, Baumgarner and Raftery (2005) have developed an iterative method that is much faster for large numbers of regressors (about 5000 in their case), but is no longer guaranteed to find the model with the highest posterior probability. This iterative BMA version has been applied in the growth context by Eicher, Papageorgiou and Roehn (forthcoming).

formative" or "reference" priors, but there is debate about the extent to which a prior can be totally noninformative, and so we use the term "default prior" here. Priors on parameters may affect results since they may influence the integrated likelihood (2), which is a key component of the posterior weights used in the averaging process (3). The integrated likelihood of a model is approximately proportional to the prior density of the model parameter evaluated at the posterior mode (Kass and Raftery, 1995). Thus the prior density should be spread out enough so that it is reasonably flat over the region of the parameter space where the likelihood is substantial. It is crucial to note, however, that the prior density should be no more spread out than necessary, since increasing the spread of the prior tends to decrease the prior ordinate at the posterior model, which decreases the integrated likelihood and may unnecessarily penalize larger models (Raftery, 1996). Thus there is a trade-off, and the various priors we discuss below make this trade-off in different ways.

We focus on a set of 12 candidate default priors that have been prominently advocated in the literature. Table 1 presents, describes and provides sources for all 12 priors. First is the *Unit Information Prior* (UIP), which contains about the same amount of information as a typical single observation (Kass and Wasserman 1995; Raftery, 1995). Second, the *Data-Dependent* prior suggested by Raftery, Madigan and Hoeting (1997) is explicitly designed to be relatively flat over the region of the parameter space supported by the data but no more spread out than necessary. Third, ten *automatic* priors used in FLS (2001b) do not rely on input from the researcher or information in the data, but only on the sample size and the number of regressors.

The first prior that we consider is defined implicitly, by the form of the integrated likelihood that is used, namely,

$$\log pr(D|M_k) \approx c - \frac{1}{2} BIC_k, \tag{6}$$

where

$$BIC_k = n \log(1 - R_k^2) + p_k \log(n). \tag{7}$$

In (7), R_k^2 and p_k are the R^2 value and the number of regressors, respectively, for model M_k, and c is a constant that does not vary across models and so cancels in the model averaging. BIC_k is the Bayesian Information Criterion for M_k, which is equivalent to the approximation derived by Schwarz (1978). The approximate integrated likelihood in (6) was the basis of the model averaging method of Raftery (1995) for linear regression, and was also used by SDM.

It follows from the results of Kass and Wasserman (1995) that for any pairwise model comparison, the ratio of posterior model probabilities resulting from the use of (6) closely approximates the ratio of posterior model probabilities that would be obtained from a particular prior for the regression parameters, with a relatively small error that is of the order $O\left(n^{-1/2}\right)$. This is a multivariate normal prior centered at zero with variance matrix equal to n times the inverse Fisher information matrix. This prior is much more spread out than the likelihood, and typically is relatively flat where the likelihood is substantial (Raftery, 1999). It contains the same amount of information as would be contained on average in a single observation and so, following Kass and Wasserman (1995), we call it the *Unit Information Prior* (UIP). Because of its simplicity and intuitive appeal, we use this UIP as a baseline, and we compare other proposed default priors to it.

Next we consider ten *automatic* priors suggested by FLS (2001b) to be applied in situations when the researcher has little or no subjective prior information. These parameter priors are based on Zellner (1986), who suggested a particular form of the *natural conjugate* gamma family of priors, namely a g-prior density for the parameters in (1):

$$p(\sigma|M_k) \propto 1/\sigma, \tag{8a}$$

$$p(\alpha|M_k) \propto 1, \tag{8b}$$

$$\beta^{(k)}|\sigma,\, M_k \sim N\left(0,\, \sigma^2\left(g_k Z^{(k)\prime} Z^{(k)}\right)^{-1}\right), \tag{8c}$$

where $Z^{(k)}$ is the $n \times p_k$ matrix consisting of the p_k regressors included in M_k, each one centered by subtracting its mean. These are all special cases of Zellner's (1986) g-prior for β_k where the g value is a factor of proportionality that scales the reciprocal of the variance of the parameter prior. Values of g that are closer to zero imply priors that are less informative, and $g = 1$ implies that prior information and data information are weighted equally in the posterior distribution.

Different automatic priors result from different choices of g_k, as listed in Table 1. The choice $g = 1/n$ (Prior 12 in Table 1) is in the spirit of the UIP. Alternatives are Prior 4, $g = \sqrt{1/n}$, which attributes a smaller asymptotic penalty than BIC, and Prior 2, $g_k = p_k/n$, where prior information increases with the number of regressors in the model. Other priors suggested by FLS (2001b) correspond to previous proposals: Priors 6 and 7 in Table 1 are versions of the Hannan and Quinn criterion (Hannan and Quinn, 1979), and Prior 9, $g_k = 1/p_k^2$, corresponds to the Risk Inflation Criterion (RIC) of Foster and George (1994), designed to take account of the number of candidate regressors. FLS (2001b) stated that Prior 8 in Table 1 is comparable to model averaging using AIC-based weights. Prior 10 is the preferred prior of FLS (2001b), developed on the basis of their experiments with their priors. It is composed of either the RIC-based prior (Prior 9) or Prior 12, depending on the number of observations and regressors in the particular dataset. For the datasets considered in this paper, Prior 10 is identical to Prior 9.

An alternative class of *data-dependent* priors can be viewed as approximating the subjective prior of an experienced researcher. Clearly, if such knowledge is readily available, it should be introduced into the analysis, and Wasserman (2000) showed that data-dependent priors can improve predictive performance.[7] Raftery, Madigan and Hoeting (1997) automate a process that specifies data-dependent priors that are as concentrated as possible, subject to being reasonably flat over the region of parameter space where the likelihood is not negligible. Their prior (Table 1, Prior 11) is determined by four hyperparameters that are explained in Table 1. A variant of such data-dependent priors is based on Laud and Ibrahim (1996) (Table 1, Prior 8) who specified $g = \delta\gamma^{1/p_j} \Big/ \left(1 - \delta\gamma^{1/p_j}\right)$. Given FLS's suggestions for γ and δ, they mention that model comparisons based on the resulting log integrated likelihood can roughly be compared to those based on the *Akaike Information Criterion* (AIC) (Akaike, 1974).

2.3 Model Priors

We consider two main types of prior over the model space that have been widely advocated. The first is the uniform prior, that gives equal prior probabilities to each model, so that $pr(M_k) = 1/K$ for each k. This seems to have been suggested first by Raftery (1988) and, for linear regression models, by George and McCulloch (1993). Hoeting et al. (1999) cite the extensive evidence that supports the good performance of the uniform model prior, since the integrated likelihood on the model space is often concentrated enough for the results to be insensitive to moderate deviations from the uniform prior.

Treating all models equally a priori might be a "neutral" choice, but it may not be the best choice when prior information is available. "Subjective Bayesianism" is prominent in economic applications, where researchers hold beliefs about the true theory that are strong enough to suggest model sizes that deviate significantly from uniform distributions. For example, in the growth context SDM argue that most researchers prefer small model sizes and suggest a prior model size

[7] FLS (2001b) point out a common criticism of data-dependent priors, namely that the posterior distribution can no longer be interpreted as a conditional distribution given the observables.

of 7 regressors. This argument is made although over 140 candidate regressors have been proposed for growth regressions in a survey for the Handbook of Economic Growth (Durlauf, Johnson and Temple, 2005). At the heart of SDM's argument is the observation that several new growth theories have been developed over the past decade. Under uniform model priors such an increase in candidate theories (along with the associated increase in candidate regressors) would imply larger expected prior model sizes although there is no reason to suspect that the true model size has changed.

Mitchell and Beauchamp (1988) suggested assigning a discrete prior probability mass to any regressor that is to be considered for exclusion from regression model M_k, namely

$$pr(M_k) = \prod_{j=1}^{p} \pi_j^{\delta_{kj}} (1 - \pi_j)^{1-\delta_{kj}}, \tag{8}$$

where $\delta_{kj} = 1$ if X_j is included in M_k and 0 otherwise. In (8), π_j is the prior probability that X_j is included in the model, and often the π_j's are equal, with $\pi_j = \pi$. This prior over model space has been widely used, for example by George and McCulloch (1993), Madigan and Raftery (1994) and SDM. Letting $\bar{p} = p\pi$ be the prior expected number of regressors in the model, SDM argued for the use of $\bar{p} = 7$ in growth applications. Ley and Steel (2007b) evaluated that choice relative to Prior 9. We evaluate the SDM choice relative to other popular parameter priors and choices of $\bar{p}$ in our experiments below. Note that when $\pi = 0.5$, the prior in (8) reduces to the uniform prior, in which case $\bar{p} = p/2$. Brown, Vannucci, and Fearn (1998; 2002) and Ley and Steel (2007b) went one step further and suggested that the probability that a given variable is included in a model itself be a random variable drawn from some distribution.

George (1999) pointed out that if several candidate regressors are highly correlated, independent model priors such as the uniform prior and the SDM prior can give too little weight to models that exclude all these variables, and he proposed a dilution prior to deal with this. A clear example of this is when three different but highly correlated measures of the same quantity are used (say three different measures of unemployment). Then with the uniform prior, the prior probability of unemployment having an effect would be 7/8, not 1/2. George's dilution prior increases the prior probabilities of models not containing the correlated regressors to take account of this effect. This seems reasonable when the variables are indeed measures of the same thing, as in the example just mentioned. However, often in economics, variables represent distinct concepts, but are nevertheless correlated. In this situation, independent prior inclusion probabilities (as in either the uniform prior or the SDM prior) do seem defensible. A direct use of the correlation-based dilution prior of

George (2001) for our growth dataset would lead to very harsh penalization of larger models.

A related situation is when several correlated regressors are proxies for the same theory. For this situation, Durlauf, Kourtellos and Tan (2006; 2007) proposed a modification of George's (2001) correlation-based dilution prior in which dilution takes place "in blocks," within the regressors that proxy each growth theory considered. This seems like a good idea in principle. However, one difficulty with it in the present context is that it requires agreement on which theories are represented by which variables. Such agreement is often not present, and in its absence a dilution prior based on the assessments of one research group cannot be viewed as a default prior. Also, for our growth dataset it would tend to strongly favor models in which each theory was represented by just one proxy regressor, so the decision about how to group the variables into blocks could be highly consequent. As a result, we have considered only independent model priors in the present paper.

2.4 An Integrated Approach to Prior Distributions and Posterior Inference

The examination of the impact of alternative model and parameter prior specifications is hampered by the diverse set of priors and the large variation in software packages used for implementation. Our baseline UIP is contained in Raftery's (1995) BICREG function written in the open source statistical language R and the data-dependent prior is contained in the MC3.REG function also written in R; these are both included in the R package BMA available at http://cran.r-project.org (Raftery, Painter and Volinsky, 2005). Priors 2-7, 9, 10 and 12 are considered in FLS's (2001b) procedure written in FORTRAN77. All of the above mentioned packages assume a uniform prior over the model space. Alternative model priors are considered in SDM's procedure available in GAUSS, but only for Prior 1. In this paper we apply one integrated procedure, programmed in R to examine the BMA results using a popular growth dataset and simulated datasets.[8]

3 Uncovering Robust Growth Determinants

Since economic growth is the fundamental driver of living standards, it is of great interest to economists and policymakers alike to identify which of the numerous proposed theories receive support from the data and which determinants are related to growth. Attempts to identify robust growth determinants date back to the extreme bound analysis of Levine and Renelt (1992) and

[8] The R program BMA.COMPARE simultaneously evaluates all 12 different parameter priors and any specific prior model size, as well as their predictive performance. It is available upon request from the first author.

Sala-i-Martin (1997). Formal BMA analysis was conducted by FLS (2001a) and SDM (2004). The dataset used across studies always contains a core of at least 41 candidate regressors, motivated by Sala-i-Martin (1997) and FLS (2001a). We base our growth analysis on this same dataset that FLS kindly shared with us.

3.1 Effects of Parameter Priors on Growth Determinants

For datasets with small numbers of observations, priors may well be expected to play an important role unless the data contain decisive information. Given that our growth dataset has 72 observations, priors may be suspected to influence growth results. As can be seen in Figure 1, the precisions of the parameter priors vary widely; for example the information contained in Prior 7 is three orders of magnitude greater than that in the FLS-preferred prior. It thus seems possible that the BMA results would vary considerably between priors.

Table 2 reports the BMA posterior inclusion probabilities for all 12 prior distributions applied to the growth dataset. Jeffreys (1961) proposed rules of thumb, refined by Kass and Raftery (1995), suggesting that the evidence for a regressor having an effect is either *weak*, *positive, strong*, or *decisive* when the posterior inclusion probabilities range from 50-75%, 75-95%, 95-99%, and > 99%, respectively.[9] We mark all variables that exhibit evidence for an effect (above 50%) with a shaded box in Table 2.[10]

Posterior inclusion probabilities and the number of regressors that exhibit evidence of an effect on growth vary widely across priors. The latter ranges from a low of 7 regressors (Priors 5, 7, and 11) to a high of 21 regressors or more (Priors 1, 3, and 12). Recall that the prior distributions are all centered at zero and that Priors 5, 7, and 11 have small prior variance (Figure 1). Priors 5, 7, and 11 thus contain strong information *against* a large effect, and the information contained in the data is too weak to overwhelm that prior. As the priors over the parameter space become sufficiently spread out to include regions where the likelihood is substantial, the number of regressors that exhibit an effect increases. However, once the priors become very spread out (especially for Prior 9), we observe a decline in the integrated likelihood, and the number of parameters that show an effect is reduced. The relationship between prior variance and the number of regressors exhibiting an effect is plotted in Figure 1.

[9] In economics an alternative rule of thumb for an effect was suggested by SDM that the posterior inclusion probability exceeds the prior inclusion probability. We discuss this rule in Section 3.2 below.

[10] Barbieri and Berger (2004) show that under regularity assumptions the median probability model (i.e., the model containing all covariates whose marginal posterior inclusion probability exceeds 50%) minimizes predictive squared error loss.

Figure 2 shows scatterplots of posterior inclusion probabilities generated by the various priors against our baseline prior (Prior 1). Since Prior 1 was the most optimistic, with 22 candidate regressors showing an effect in Table 2, it is no surprise that most of the points in the scatterplots lie above the 45 degree line, indicating generally higher posterior inclusion probabilities for each regressor under Prior 1 as compared to other priors. More importantly, however, the scatterplots highlight not only that Prior 1 is more optimistic, but also how the differences between Prior 1 and alternative priors increase as the implied g-prior diverges. Priors 1, 6, and 12 have relatively similar results, but most other priors show differing effects implied by the priors.

Alternatively, one might be tempted to interpret Table 2 as suggesting that 6 regressors (Confucius, Initial GDP, Life Expectancy, Rule of Law, Sub-Saharan Africa dummy, and Equipment Investment) are robustly related to growth, since there is clear evidence for an effect for each of these regressors across all priors. We view this interpretation as misguided because the selection criterion based on the lowest common denominator is inappropriately conservative. Instead we argue that the choice of variable selection method should be based on comparing the predictive performances of the prior distributions.

The same dispersion of results across priors can be observed in the number of regressors contained in the models that have the greatest posterior model probability. Table 3 reports the best model for each prior and shows that the best model discovered by Prior 1 also generates the highest adjusted R^2. Image plots in Figures 3a and 3b compare Prior 1 with the FLS benchmark prior (Prior 9) using the growth dataset. The figures highlight models used in the averaging process (ranked by posterior probability on the vertical axis, where the vertical width indicates the model weight). Red and blue indicate the inclusion of a regressor with a positive and negative coefficient, respectively. The image plots highlight how different the models are over which the various priors average. It is no surprise that these two prominent priors differ substantially in terms of posterior inclusion probabilities.

Economists are interested not just in *which variable* exhibits evidence of an effect, but also in *what economic impact* a variable has. Table 4 shows that the posterior inferences about the effects of variables are much more robust to the prior specification than are inferences about the best model. This table presents the posterior mean and standard deviation of the regression coefficients for all 12 priors, and indicates that the economic impact of regressors, as measured by the posterior mean of the corresponding coefficients, hardly varies across priors. With the exception of Equipment Investment and High School Enrollment, we find that economic significance is estimated quite uniformly. This indicates that the estimated economic effects of most variables are relatively robust to the prior specification, although the models selected may vary significantly across prior parameter distributions. Figure 4 shows all standardized posterior means for each candidate regressor, which indicate general agreement on the economic impact. None of the posterior means have reversed signs and they are generally of similar magnitudes.

Table 5 shows that our results do not depend on the computational algorithms used, which generate similar inclusion probabilities. However, the Raftery (1995) branch-and-bound algorithm is faster and it also discovers the best model, unlike the other algorithms. We also include an example of the class of general-to-simple approaches based on multi-path search proposed in Hoover and Perez (1999). The PcGets algorithm (Hendry and Krolzig, 2004) is shown for comparison as it has been suggested as an alternative to BMA model selection (see Durlauf et al. 2006, 2007). Table 5 shows that the best model identified by BMA with the branch-and-bound algorithm is better than the one selected by PcGets in terms of both adjusted R^2 and BIC. More fundamentally, the PcGets approach does not incorporate uncertainty about model form, unlike the BMA approaches. Carrying out inference conditional on the selected PcGets model could substantially underestimate uncertainty, as dramatically illustrated by Freedman (1983) in the context of a similar general-to-simple algorithm for selecting predictor variables in regression. Although we have not assessed the predictive performance of the PcGets-selected model here, the theoretical and empirical results summarized by Raftery and Zheng (2003) indicate that BMA has better predictive performance than approaches based on single selected regression models.

3.2 Combined Effects of Parameter and Model Priors on Growth Determinants

In Bayesian analysis, any valuable prior knowledge should be included in the priors. This "subjective Bayesian" approach has become prominent in economics applications, especially in the field of economic growth, where SDM argued that uniform priors on the model size are not desirable in this context. Instead, SDM proposed that the true growth model should be closer to 7 regressors.

In addition, SDM contended that their alternative prior distribution also requires a new effect-threshold to identify candidate regressors that exhibit an effect on the dependent variable. Their suggested effect-threshold is that the posterior regressor probability must exceed the prior model probability. In our growth dataset the SDM benchmark model prior implies that 7 of 41 regressors matter to the analysis, which yields an effect-threshold of each individual regressor of $7/41 = 17\%$. However, there seems to be a tension between this and the basic idea of Bayesian statistics, namely that all information about a quantity of interest is contained in the posterior distribution, which is determined by the prior and the likelihood. The tension arises when effect-thresholds are based only on prior information. When the posterior inclusion probability of a regressor is below 50%, the evidence for it is weak, *even if* prior evidence would lead one to expect that the model should be small. We favor the conventional Jeffreys rules of thumb, in particular the 50% threshold for reporting effects, since it implies that the *combined* available evidence (of prior and likelihood) does not support the variable having an effect when the posterior inclusion probability of a variable falls below 50%.

The SDM effect-threshold has two important implications. First, the smaller the number of regressors specified by the model prior, the *lower* the threshold on the posterior inclusion probability of an individual regressor. So on the one hand the researcher imposes priors that favor smaller models, but on the other hand the effect-threshold, in terms of inclusion probabilities of the individual variables, is lowered. Researchers that stipulate strong priors over small models are at the same time relaxing the effect-threshold. Second, as highlighted by the example of growth theories, as the number of candidate regressors rise, but the prior model size stays constant, the effect-threshold becomes lower. We discuss results with both the Jeffreys and SDM effect-thresholds.

Table 6 shows how the results differ between the two kinds of model prior. As expected, the subjective prior expectation that the true growth model contains only 7 covariates leads to smaller models than the uniform model prior, ranging from 3 to 8 effective regressors for the Jeffreys effect-threshold and from 6 to 12 effective regressors for the SDM effect-threshold. Again the priors with intermediate variance have a slightly larger number of regressors (Priors 3, 4, and 12), and as before the number of regressors that exhibit an effect declines as the prior variance become large (Priors 6 and 9). One change is that using the SDM threshold some new regressors, such as Muslim, Years Open, and Protestant, become important for a number of priors. This leads not only to fewer regressors that surpass the effect-threshold, but also to a *different* set of effective regressors.

The restrictive model prior has the least impact on Prior 11; for this prior, the Rule of Law variable loses significance but otherwise the results are identical to Table 2. Thus forcing BMA to increase the weight on smaller models and penalize larger models affects priors differently: it can change the number of candidate regressors that pass the effect-threshold, and it can lead to different regressors with high inclusion probabilities. Thus, not only does the nonuniform model prior lead to smaller models, but it also attributes significance to regressors that were previously not seen as strong. Later we examine these results in light of predictive performance to assess whether the researchers' subjective prior was indeed appropriate.

The two different effect-thresholds do not alter the results dramatically. As expected, the more stringent Jeffreys threshold is more limiting in terms of the number of regressors that show an effect. The weaker threshold adds between two to four regressors to the list of effective growth determinants. For the sake of clarity and to establish constant and unambiguous thresholds, we suggest the traditional Jeffreys' thresholds as a default.

The image plot in Figure 3c shows Prior 1 with prior expected model size 7, the combination recommended by SDM. Comparing Figures 3a and 3c, we see that the model prior with prior expected model size 7 biases the results towards growth models with fewer variables. We will see in the subsequent section on predictive performance that this does not improve prediction. Note also the similarity between Figures 3b and 3c, two very different model and parameter priors. This similarity was first observed for these specific priors by Masanjala and Papageorgiou (2005). Ley and Steel (2007b) describe the similarity between the FLS uniform prior and Prior 1 with prior model size 7 as arising "mostly by accident" and discuss specific parameter constellations that generate similar posterior probabilities.

This similarity has a theoretical explanation. Any two prior structures may differ according to the parameter variance (proportional to $1/g$) and the model prior size, $\bar{p}$. Comparing the posterior probabilities for a given model in (3) for different prior structures, Kass and Raftery (1995) show that an increase in the prior standard deviation by a factor c, is approximately equivalent to a reduction in the prior odds for an increase in the model size by an additional variable, by the same factor of c.

Using the approximation of Kass and Raftery (1995, equation 14), it can be shown that for two prior structures, A, B, with associated prior scale factors, g_A, g_B, and expected prior model sizes, $\bar{p}_A$, $\bar{p}_B$, the posterior odds for one regression model against an alternative regression model with one additional regressor are approximately equal when the prior structures satisfy

$$\frac{g_A}{g_B} = \left[\frac{\bar{p}_B}{\bar{p}_A}\frac{(p-\bar{p}_A)}{(p-\bar{p}_B)}\right]^2 . \qquad (9)$$

For the FLS dataset with $n = 72$ and $p = 41$, the FLS benchmark parameter prior implies $g_A = 1/p^2$, combined with the uniform model prior, $\bar{p}_A = p/2$. When $g_B = 1/n$ as in the case of Prior 1, used by SDM, equation (9) holds when the prior expected model size is $\bar{p}_A = 7.03$. It is therefore not surprising that for the SDM suggested prior expected model size of $\bar{p}_A = 7$, the priors recommended by SDM and FLS yield similar results for the growth dataset, although they are based on very different parameter and model priors. Note that this similarity depends crucially on the number of parameters in the dataset, p. For the values of g suggested by SDM and FLS, the prior expected model size that equates posterior probabilities, $\bar{p}_A$, therefore depends on the number of observations. This indicates that there is a tradeoff between prior expected model size and prior parameter variance. Subjective priors that favor small models thus achieve their aim by punishing larger models (Figure 3c) or by increasing the prior variance on each individual parameter (Figure 3b).

In summary, candidate default priors differed considerably in dispersion, and led to the choice of different sets of variables. As few as 3 and as many as 22 regressors were found to be related to growth, depending on the specific parameter and/or model prior used. In contrast, the BMA posterior effect estimates and standard errors were quite robust to the prior specification.

3.3 Assessment of Prior Distributions using Predictive Performance

The previous analysis does not identify the best prior for our growth dataset. Instead, we compare the priors on the basis of their predictive performance. Prediction provides a neutral criterion to compare methods. To assess predictive performance, we outline the scoring rules used compare the performances of the different methods, and we assess the impact of both parameter and model priors on predictive performances for the growth and simulated datasets.

We base our predictive performance evaluation on three different scoring rules: the Mean Squared Error (MSE), the Continuous Ranked Probability Score (CRPS; Matheson and Winkler, 1976), and the Log Predictive Score (LPS; Good, 1952). The CRPS has been widely used in other areas such as weather forecasting, but this is its first use in economics of which we are aware. Scoring rules provide summary measures to evaluate probabilistic forecasts; they assign a numerical score based on the value that materializes relative to the forecast. All three are proper scoring rules for assessing predictive performance.[11]

The MSE is the most popular measure to assess predictive performance in economics. It focuses on point estimation, while the LPS and the CRPS assess the entire predictive distribution. The CRPS and the LPS assess both the sharpness of a predictive distribution and its calibration, namely the consistency between the distributional forecasts and the observations. However, the LPS assigns harsh penalties to particularly poor probabilistic forecasts, and can be very sensitive to outliers

[11] A proper scoring rule is one in which the forecaster gets the best score by reporting a forecast distribution that mirrors his or her true beliefs. The scoring rule is strictly proper if its maximum is unique.

and extreme events (Weigend and Shi, 2000; Gneiting and Raftery, 2007). This may be a factor when we split our small sample to examine predictive performance. The CRPS is more robust to outliers (Carney, Cunningham and Byrne, 2006; Gneiting and Raftery, 2007), and hence it is our preferred measure of the performance of the predictive distribution as a whole. We also report the LPS for comparability with previous work, notably that of FLS (2001b). The three scoring rules are described in the appendix.

Formally, the predictive distribution in BMA is as follows. Let q be a quantity of interest, such as an out of sample observation. Then the posterior distribution of q, given the data D, is given by

$$pr\left(q|D\right) = \sum_{k=1}^{K} pr\left(q|M_k, D\right) pr\left(M_k, D\right), \tag{10}$$

which is the average of the posterior distributions of q under the different models, weighted by the posterior model probabilities. In our application, the predictive ability of a prior distribution is evaluated and compared to the predictive ability of alternative prior distributions.

The analysis requires us to split the sample into a training set, D^T, and a hold-out set, D^H. The training sample is used to derive the BMA results, and the hold-out sample allows us to gauge the predictive performance of on independent data. The method of cross validation via training and hold out sets dates back at least as far as Mosteller and Wallace (1963). Our split of the data involves a training set that contains 80% of the data and thus leaves 20% of the data to be predicted. The larger the dataset, the more desirable even splits become.

Each random split of the data generates a different fit and therefore a different score. Given the different functions that each scoring rule optimizes, it is not surprising that they may not agree on which particular prior generates the highest score for any given data split. We therefore used S random splits rather than a single one, and found that we needed $S \geq 200$ to obtain reliable results. Since we are comparing each prior to our baseline prior, we also report the proportion of times that Prior 1 outperformed the prior being evaluated.

Table 7 shows the predictive performance of the 12 parameter priors in conjunction with uniform model priors as evaluated by the MSE, LPS and CRPS using $S = 578$ random splits. The MSE and the CRPS agree that our baseline Prior 1 decisively outperforms all the other priors. The LPS suggests, however, that Priors 2, 4, 6, and 8 outperform Prior 1. Since this result runs counter to the results from the two other scoring rules, it seems possible that the difference is due to outliers or influential cases in the dataset. Several of the regressors have extreme outlying values. When such cases are in the test set, they can have a large effect on the LPS, while the CRPS is more robust to individual cases. Given the known outlier sensitivity of the LPS, we discount the results it gives for this dataset, and conclude that Prior 1 performs best in this case.

Next, we compare the default model priors to the more restrictive model prior proposed by SDM by assessing whether such a prior generates better predictive performance. Table 8 shows the predictive performance of a number of alternative (smaller) prior model probabilities for the growth dataset including prior model size 7 (shaded area). This shows that nonuniform model priors that favor smaller models do not provide improved predictive performance in the growth dataset. The results indicate that the UIP does not overfit in the growth application.

Indeed in the growth example, our baseline prior is shown to be flat enough to extend over the part of the distribution where the likelihood is substantial, but not so flat that it overpenalizes large models (as is for example Prior 9). This is clear from the fact that the UIP performs better than parameter and model priors that specify either smaller model sizes, or smaller or larger variances of the prior parameter distribution. Overall, the unit information prior (Prior 1) with a uniform model prior performs best of the candidate default priors that we have evaluated. Thus the prior expectation of a model size of about 7 regressors is not borne out by the predictive performance results. Indeed only as the prior model size approaches $p/2$ do some of the other priors again show better performance. Note however, that no other prior considered ever beats Prior 1 for any model size in terms of MSE or CRPS.

4 Simulated Data

In the growth dataset we found that the models selected were sensitive to the prior used, although posterior inference about effect sizes was relatively robust. We found that one candidate default prior, the Unit Information Prior, dominated the others in terms of predictive performance. The question is whether this result is specific to the growth dataset, or whether it applies more generally. To investigate this question we now apply BMA to several simulated datasets designed to mimic features of datasets commonly found in economics.

4.1 Effects of Prior Structure

We examine the effects of the set of priors using simulated datasets from two models that have been prominent in the BMA literature: Model 1 is provided by FLS and is based on Raftery, Madigan and Hoeting (1997), and Model 2 was also suggested by FLS and is based on George and McCulloch

(1993). For Model 1 we generate an $n \times p\,(p = 15)$ matrix $R = (r_1, ..., r_{15})$ of regressors, where the first ten columns are drawn from independent standard normal distributions, and the next five columns are constructed according to $(r_{11}, ..., r_{15}) = (r_1, ..., r_5)\,(0.3\ 0.5\ 0.7\ 0.9\ 1.1)'\,(1, 1, 1, 1, 1) + E$, where E is an $n \times 5$ matrix of independent standard normal deviates. Model 1 implies small to moderate correlations between the first and last five regressors $r_1, ..., r_5$ and $r_{11}, ..., r_{15}$. The correlations increase from 0.153 to 0.561 for $r_1, ..., r_5$ and are somewhat larger between the last five regressors, reaching 0.740. Each regressor is centered by subtracting its mean, which results in a matrix $Z = (z_1, ..., z_{15})$. A vector of n observations is then generated according to

$$\textit{Model 1}: y = 4i_n + 2z_1 - z_5 + 1.5z_7 + z_{11} + 0.5z_{13} + \sigma\varepsilon, \tag{11}$$

where the n elements of ε are independent standard normal and $\sigma = 2.5$. In Model 1 a third of all the regressors intervene, which we view as fairly typical of some real world situations, and we examine datasets with 50 and 100 observations to stay close to the structure of our growth example.

The structure of Model 2 is closer to the growth dataset in terms of numbers of observations and numbers of regressors. It is generated using p regressors, $r_i = r_i^* + e, i = 1, ..., p$, where r_i^* and e are n-dimensional vectors of independent standard normal deviates. This induces a pairwise correlation of 0.5 between all regressors. Let Z again denote the $n \times p$ matrix of centered regressors, and generate the n observations according to

$$\textit{Model 2}: y = i_n + \sum_{h=1}^{p/2} z_{(p/2+h)} + \sigma\varepsilon, \tag{12}$$

where the n elements of the error are again independent standard normal and $\sigma = 2$. In this simulation model, the second half of the regressors intervene, namely $(z_{21}, ..., z_{40})$.

For Model 1, the differences in the prior variances shown in Figures 5a,b are similar to the magnitudes observed for the growth dataset in Figure 1. Again about three orders of magnitude separate the most concentrated and most diffuse priors, although the level of concentration is a bit lower in the simulated datasets. Tables 9a,b show, however, that with well-behaved data all priors basically agree upon which regressors have an effect, even in a dataset that contains only 50 observations. For the larger simulated dataset in Model 2, with about three times the number of candidate regressors as in Model 1, we again find diversity in the number of regressors identified as having an effect on the dependent variable. Table 7c shows that several priors are clearly too concentrated, with Priors 2, 5, and 7 identifying only between 3 and 7 of the 20 relevant regressors that in fact had an effect on the dependent variable. As the prior variance increases enough to cover the more substantive part of the likelihood, the priors are able to pick up more of the relevant regressors, getting closer to the correct number of regressors. Priors 3, 9, and 11 pick up 16 candidate regressors although only Prior 1 shows appropriately high posterior inclusion probabilities.

In summary, our simulation experiment shows that priors can matter, especially when there are many candidate regressors. The Unit Information Prior is the only one that was robust across simulations, coming closest to identifying the right regressors in all cases.

4.2 Prior Structure and Predictive Performance

We now report the predictive performance results from the simulated data experiments. Table 10 shows the Unit Information Prior's overall superior performance, and also the importance of examining alternative scoring rules. In terms of point estimates, the MSE was quite consistent in its evaluation. However, the CRPS and LPS differed in their assessment at times, even after 400 trials. Prior 1 usually outperformed all the other priors for all scoring rules, but some other priors gave better CRPS results for some simulated datasets.

While the MSE and the LPS unanimously attributed the best predictive performance to Prior 1, for Model 1 the CRPS did not agree and identified Priors 3, 4, and 8 as best. For Model 2, however, the results again show strong overall support for Prior 1. The CRPS and LPS did not agree on one prior each, but otherwise there is clear evidence that Prior 1 is not overfitting. This result is not surprising since our baseline prior's number of regressors with an effect was the closest to the true number of regressors in the model.

5 Conclusion

Model uncertainty is intrinsic in economic analysis and the economic growth literature has been a showcase for model uncertainty over the past decade. Over 140 growth determinants have been motivated by the empirical literature, and the number of competing theories has grown dramatically since the advent of the New Growth Theory (see Durlauf, Johnson and Temple, 2005 for a survey). Standard in all empirical studies is, or should be, the examination of robustness and performance of alternative specifications. Bayesian Model Averaging (BMA) provides a solid theoretical foundation for robustness analysis that juxtaposes different theories. For a well-known growth dataset, we show that growth determinants can be sensitive to the prior specification. The same analysis also shows, however, the important result that model-averaged inference about the economic effect of growth regressors is robust across alternative priors.

To identify the best prior for our growth dataset, we examine the predictive performance of 12 candidate default parameter priors that have been proposed in the economics and statistics literature, as well as two candidate model priors. We argue that predictive performance is a neutral criterion for comparing different priors, and we introduce an improved scoring rule. In addition, we examine these priors' success in identifying the right determinants in simulated datasets. The Unit Information Prior (UIP) for the parameters performed consistently better than the other 11 priors in the growth data, and in simulated data, and as measured by all three scoring rules. The uniform model prior together with the uniform model prior also performed better than the Mitchell-Beauchamp model prior with expected model size 7, which had previously been recommended by Sala-i-Martin, Doppelhofer and Miller (2004) in the context of economic growth. We view the Unit Information Prior with the uniform model prior as a reasonable default prior and starting place, but our results also highlight that researchers should also assess other possibilities that may be more appropriate for their data.

In spite of widespread doubts about the ability of the "small" cross-country growth dataset to provide a rich set of growth determinants, our analysis shows that the UIP parameter prior with uniform model priors robustly identifies far more growth determinants than other priors. The UIP discovers substantial evidence for 14 additional growth determinants as compared to those in Sala-i-Martin, Doppelhofer and Miller (2004) and Fernández, Ley and Steel (2001b). Hence we show that the appropriate prior in the growth context delivers a rich set of robust growth determinants that also generate good predictive performance. The new regressors prominently feature colonial origins, openness (Outward Orientation as well as the Black Market Premium) as well as institutional characteristics (Rule of Law, Civil Liberties and Ethnolinguistic Fragmentation). Thus our results provide support for several new growth theories.

References

Akaike, H. (1974). "A new look at the statistical model identification". *IEEE Transactions on Automatic Control* 19, 716–723.

Barbieri, M.M. and J.O. Berger. (2004). "Optimal Predictive Model Selection," *Annals of Statistics* 32, 870-897.

Bates, J.M. and C.W.J. Granger. (1969). "The Combination of Forecasts," *Operations Research* 20, 451-468.

Begun, J. and T.S. Eicher. (2006). "In Search of a Sulphur Dioxide Environmental Kuznets Curve: A Bayesian Model Averaging Approach," working paper, University of Washington.

Brier, G. (1950). "Verification of Forecasts Expressed in Terms of Probability," *Monthly Weather Review* 78, 1-3.

Brock, W. and S.N. Durlauf. (2001). "Growth Empirics and Reality," *The World Bank Economic Review* 15, 229-272.

Brock, W., S.N. Durlauf and K. West. (2003). "Policy Evaluation in Uncertain Economic Environments," *Brookings Papers on Economic Activity* 1, 235-322.

Brown, P.J., M. Vannucci, and T. Fearn. (1998). "Multivariate Bayesian Variable Selection and Prediction," *Journal of the Royal Statistical Society. Series B* 60, 627-641.

Brown, P.J., M. Vannucci, and T. Fearn. (2002). "Bayes Model Averaging with Selection of Regressors," *Journal of the Royal Statistical Society, Series B* 64, 519-536.

Carney, M., P. Cunningham and S. Byrne. (forthcoming) "The Benefits of Using a Complete Probability Distribution when Decision Making: An Example in Anticoagulant Drug Therapy," *Medical Decision Making.*

Clyde, M. and E.I. George. (2004). "Model Uncertainty," *Statistical Science* 19, 81-94.

Clyde, M., H. DeSimone and G. Parmigiani. (1996). "Prediction Via Orthogonalized Model Mixing," *Journal of the American Statistical Association* 91, 1197-1208.

Doppelhofer, G. (2007, forthcoming). "Model Averaging" in The New Palgrave Dictionary in Economics, 2nd edition. L. Blume and S. Durlauf (eds.).

Durlauf, S.N., P. Johnson and J. Temple. (2005). "Growth Econometrics," in *Handbook of Economic Growth*, P. Aghion and N. Durlauf, eds., North Holland, Amsterdam.

Durlauf, S.N., A. Kourtellos and C.-M. Tan. (2006). "Is God in the Details? A Reexamination of the Role of Religion in Economic Growth," working paper, University of Wisconsin.

Durlauf, S.N., A. Kourtellos and C.-M. Tan. (forthcoming). "Are Any Growth Theories Robust?" *Economic Journal.*

Efroymson, M.A. (1960). "Multiple Regression Analysis," in *Mathematical Methods for Digital Computers*, edited by A. Ralston and H. S. Wilf. Wiley, New York.

Eicher, T.S., C. Henn, and C. Papageorgiou. (2007). "Trade Creation and Diversion: Model Uncertainty, Natural Trading Partners, and Robust PTA Effects," working paper, University of Washington.

Eicher, T.S., C. Papageorgiou and O. Roehn. (2007, forthcoming). "Unraveling the Fortunes of the Fortunate: An Iterative Bayesian Model Averaging (IBMA) Approach," *Journal of Macroeconomics*.

Epstein, E. (1969). "A Scoring System for Probabilities of Ranked Categories," *Journal of Applied Meterology* 8, 985-987.

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

Fernández C., E. Ley and M.F.J. Steel. (2001b). "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics* 100, 381-427.

Foster, D.P. and E.I. George. (1994). "The Risk Inflation Criterion for Multiple Regression," *The Annals of Statistics* 22, 1947-1975.

Freedman, D.A. (1983). "A Note on Screening Regression Equations," *The American Statistician* 37, 152-155.

Furnival G.M. and R.W. Wilson. (1974). "Regressions by Leaps and Bounds," *Technometrics* 16, 499-511.

Garratt, A., K. Lee, M.H. Pesaran and Y. Shin. (2003). "A Long Run Structural Macroeconometric Model of the UK," *Economic Journal* 113, 412-455.

George, E.I. (1999). "Sampling Considerations for Model Averaging and Model Search," Invited discussion of "Bayesian Model Averaging and Model Search Strategies by M.A. Clyde". *Bayesian Statistics,* 6, (J.M.Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith, eds.), Oxford University Press.

George, E.I. (2001). "Dilution Priors For Model Uncertainty," University of Texas MSRI Workshop on Nonlinear Estimation and Classification, Berkeley, California.

George, E.I. and R.E. McCulloch. (1993). "Variable Selection via Gibbs Sampling," *Journal of the American Statistical Association* 88, 881-889.

Gneiting, T., Balabdaoui, F. and Raftery, A.E. (2007). "Probabilistic Forecasts, Calibration and Sharpness," *Journal of the Royal Statistical Society, Series B* 69, 243-268.

Gneiting, T. and A.E. Raftery. (2007). "Strictly Proper Scoring Rules, Prediction and Estimation," *Journal of the American Statistical Association* 102, 359-378.

Good, I.J. (1952). "Rational Decisions," *Journal of the Royal Statistical Society, Series B* 14, 107-114.

Hannan, E.J. and B.G. Quinn. (1979). "The Determination of the Order of an Autoregression," *Journal of the Royal Statistical Society, Series B* 41, 190-195.

Hendry, D.F. and H.-M. Krolzig. (2004). "We Ran One Regression," *Oxford Bulletin of Economics and Statistics* 66, 799-810.

Hersbach, H. (2002). "Decomposition of the Continuous Ranked Probability Score for Ensembles Prediction Systems," *Weather and Forecasting* 15, 559-570.

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

Hoover, K.D., and Perez, S.J. (1999). "Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search," *Econometrics Journal* 2, 167-191.

Jackman S. and B. Western. (1994). "Bayesian Inference for Comparative Research," *American Political Science Review* 88, 412-23.

Jeffreys, H. (1961). *The Theory of Probability.* Oxford University Press. *Journal of the American Statistical Association* 91*,* 1197-1208*.*

Kass, R.E. and A.E. Raftery. (1995). "Bayes Factors," *Journal of the American Statistical Association* 90, 773-795.

Kass, R.E. and L. Wasserman. (1995). "A Reference Bayesian Test for Nested Hypotheses and its Relationship to the Schwarz Criterion," *Journal of the American Statistical Association* 90, 928-934.

Laud, P.W. and J.G. Ibrahim. (1996). "Predictive Specification of Prior Model Probabilities in Variable Selection," *Biometrika* 83, 267-274.

Leamer, E.E. (1978). *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley, New York.

Leon-Gonzalez R. and D. Montolio. (2004). "Growth, Convergence and Public Investment, A Bayesian Model Averaging Approach," *Applied Econometrics* 36, 1925-1936.

Levin, A. and J. Williams. (2003). "Robust Monetary Policy with Competing Reference Models," *Journal of Monetary Economics* 50, 945-975.

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

Ley, E. and M.F.J. Steel. (2007a forthcoming). "Jointness in Bayesian Variable Selection with Applications to Growth Regression," *Journal of Macroeconomics.*

Ley, E. and M.F.J. Steel. (2007b). "On the Effect of Prior Assumptions in BMA with Applications to Growth Regression," working paper, World Bank.

Madigan, D. and A.E. Raftery. (1994). “Model Selection and Accounting for Model Uncertainty in Graphical Models using Occam’s Window,” *Journal of the American Statistical Association* 89, 1535-1546.

Madigan, D. and J. York. (1995). “Bayesian Graphical Models for Discrete Data,” *International Statistical Review* 63, 215-232.

Malik A. and J. Temple. (2006). “The Geography of Output Volatility,” working paper, University of Bristol.

Masanjala,W.H. and C. Papageorgiou. (2005). “Initial Conditions, European Colonialism and Africa’s Growth,” working paper, Louisiana State University.

Masanjala, W.H. and C. Papageorgiou. (2007a, forthcoming). “A Rough and addressStreetLonely Road to Prosperity: A Re-examination of Sources of Growth in placeAfrica using Bayesian Model Averaging,” *Journal of Applied Econometric*s.

Masanjala, W.H. and C. Papageorgiou. (2007b). “Initial Conditions and Post-War Growth in sub-Saharan Africa,” unpublished working paper, IMF.

Matheson, J. and R. Winkler. (1976). “Scoring Rules for Continuous Probability Distributions,” *Management Science* 22, 1087-1095.

Min, C.K. and A. Zellner. (1993). “Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates,” *Journal of Econometrics* 56, 89-118.

Mitchell, T.J. and J.J. Beauchamp. (1988). “Bayesian Variable Selection in Linear Regression (with discussion),” *Journal of the American Statistical Association* 83, 1023-1036.

Mosteller, F. and D.L. Wallace. (1963). “Inference in an Authorship Problem,” *Journal of the American Statistical Association* 58, 275-309.

Moulton, B.R. (1991). “A Bayesian Approach to Regression Selection and Estimation with Application to a Price Index for Radio Services,” *Journal of Econometrics* 49, 169-193.

Newbold, P. and C.W.J. Granger. (1974). “Experience with Forecasting Univariate Time Series and Combination of Forecasts (with discussion),” *Journal of the Royal Statistical Society, Series A* 137, 131-165.

Palm, F.C. and A. Zellner. (1992). “To Combine or Not to Combine? Issues of Combining Forecasts,” *Journal of Forecasting* 11, 687-701.

Raftery, A.E. (1988). “Approximate Bayes Factors for Generalized Linear Models,” *Technical Report no. 121,* Department of Statistics, University of Washington.

Raftery, A.E. (1993). “Bayesian Model Selection in Structural Equation Models,” in *Testing Structural Equation Models* (K.A. Bollen and J.S. Long, eds.), pp. 163-180, Beverly Hills: Sage.

Raftery, A.E. (1995). "Bayesian Model Selection for Social Research," *Sociological Methodology* 25, 111-163.

Raftery, A.E. (1996). "Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models." *Biometrika* 83, 251-266.

Raftery, A.E. (1999). "Bayes Factors and BIC: Comment on Weakliem," *Sociological Methods and Research* 27, 411-427.

Raftery, A.E., D. Madigan and J.A. Hoeting. (1997). "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association* 92, 179-191.

Raftery, A.E., I. Painter, and C. Volinsky. (2005). "BMA: An R Package for Bayesian Model Averaging," *R News* 5, 2-8.

Raftery, A.E. and Zheng, Y. (2003). Discussion: "Performance of Bayesian Model Averaging," *Journal of the American Statistical Association* 98, 931-938.

Sala-i-Martin, X. (1997). "I Just Ran Two Million Regressions," *AEA Papers and Proceedings* 87, 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, 813-835.

Schwarz. G. (1978). "Estimating the Dimension of a Model," *Annals of Statistics* 6, 461-464.

Sirimaneetham, V. and J. Temple. (2006). "Macroeconomic Policy and the Distribution of Growth Rates," CEPR Discussion Papers no. 5642.

Stock, J.H. and M.W. Watson. (2006). "Forecasting with Many Predictors," in *Handbook of Economic Forecasting*, Vol 1, eds. Elliot, C.W.J. Granger and A. Timmermann. North-Holland.

Wasserman, L. (2000). "Bayesian Model Selection and Model Averaging," *Journal of Mathematical Psychology* 44, 92-107.

Weigend, A.S., and S. Shi. (2000). "Predicting Daily Probability Distributions of S&P500 Returns," *Journal of Forecasting* 19, 375-392.

Yeung, K.Y., Bumgarner, R.E. and Raftery, A.E. (2005). "Bayesian Model Averaging: Development of an Improved Multi-Class, Gene Selection and Classification Tool for Microarray Data," *Bioinformatics* 21, 2394-2402.

Zellner, A. (1986). "On Assessing Prior Distributions and Bayesian Regression Analysis with g-prior Distributions," in *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti.* Goel, P.K. and A. Zellner, eds., North-Holland, Amsterdam.

Appendix

Scoring Rules

The goal in evaluating predictive performance is to maximize the *sharpness* of predictive distributions subject to *calibration* (Gneiting et al., 2007). Calibration refers to the statistical consistency between the distributional forecasts and the observations; it is a joint property of the forecasts and the values that materialize. Sharpness refers to the concentration of the predictive distributions around the observation and it is a property of the forecasts only.

Mean Square Error (MSE)

The most basic measure predictive performance is the mean squared error that focuses on point estimation and point prediction. For point estimates this is a straightforward process that involves determining the Euclidean distance between the predicted and observed points. Given the BMA posterior distribution, $pr\,(\theta|D) = \sum_{k=1}^{K} pr\,(\theta|M_k, D)\, pr\,(M_k, D)$, and the point estimate provided by the posterior mean, $E\,[\beta|D] = \sum_{k=0}^{K} \hat{\beta}\ pr\,(M_k|D)$, the MSE is defined as

$$MSE\,(\theta') = \frac{\sum_{k=1}^{K} (\theta_k - \theta)}{K}. \tag{A1}$$

Log Predictive Score (LPS)

The MSE is sufficient when researchers are concerned only with the quality of a point forecast. However, researchers might also be interested in providing a good prediction of the density, to tests whether the model produces estimates that give both, a high density at the observation and correct probability estimates. The LPS is the logarithmic scoring rule where each event, A, is assigned a score of $-\log\,[pr\,(A)]$ (Good 1952).

The predictive ability of any model is then measured by the sum of the logarithm of the posterior predictive ordinates for the observations in the hold-out set. This is the logarithm of the geometric mean of the conditional predictive ordinates. The log score for any given model is the observed coordinate of the predictive density

$$-\sum_{\theta' \varepsilon D^H} \log pr\left(\theta'|M_{k,}D^T\right), \tag{A2}$$

where $pr\left(\theta'|M_{k,}D^T\right)$ is the posterior predictive ordinate. The predictive log score for BMA is then

$$LPS\,(\theta') = -\sum_{\theta' \varepsilon D^H} \log\left\{\sum_{k=1}^{K} pr\left(\theta'|M_{k,}D^T\right) pr\left(M|D^T\right)\right\}. \tag{A3}$$

Continuous Ranked Probability Score (CRPS)

The Continuous Ranked Probability Score (CRPS; Mattheson and Winkler, 1976), is a verification method for probabilistic forecasts of continuous variables. It is equivalent to the Brier Score (Brier, 1950) integrated over all possible values and is a generalization of the Ranked Probability Score (Epstein, 1969) that is used to evaluate probabilistic predictions over ordinal variables. In essence, the CRPS measures the difference between the predicted and the occurred cumulative distributions. The squared errors are computed with respect to the cumulative probabilities of the forecast and observation

$$CRPS\left(\theta'\right) = \int_{-\infty}^{\infty} \left(pr\left(\theta'|M_{k,}D^{T}\right) - 1\left\{\theta' \geq x|\right\}\right)^{2} d\theta', \tag{A4}$$

where $1\left\{\theta' \geq x|\right\}$ denotes a Heaviside step function that attains the value 1 if $\theta' \geq x$ and the value 0 otherwise. The CRPS thus measures the area between the observed value and the predicted cumulative probability density function. Therefore *sharpness* (small spread) is rewarded if the prediction is accurate. A perfect CRPS score is 0.

Like the LPS, the mean CRPS is calculated over all predictions to determine the average error. Hersbach (2002) shows that the CRPS reduces to the MSE for deterministic forecasts. Therefore, this evaluation technique is the preferred means of comparing deterministic and probabilistic forecasting methods.

Table 1: Parameter Prior Structures

Prior	Specification of g-prior	Comment	Source
1	Unit Information Prior	The prior contains information approximately equal to that contained in a single typical observation. The resulting posterior model probabilities are closely approximated by the Schwarz Criterion, BIC.	Kass and Wasserman (1995)
2	$g_k = p_k / n$	Prior information increases with the number of regressors in the model.	FLS(2001b)
3	$g_k = p^{1/p_k} / n$	Prior information decreases with the number of regressors in the model.	FLS(2001b)
4	$g = 1/\sqrt{n}$	This is an intermediate case of prior 1 suggested by FLS where a smaller asymptotic penalty is chosen for larger models.	FLS(2001b)
5	$g_k = \sqrt{p_k / n}$	This is an intermediate case of prior 2, suggested by FLS, where prior information increases with the number of regressors in the model.	FLS(2001b)
6	$g = 1/(\ln n)^3$	The Hannan-Quinn criterion. CHQ=3 as n becomes large.	Hannan-Quinn (1979)
7	$g_k = \ln(p_k + 1)/(\ln n)$	Prior information decreases even slower with sample size and there is asymptotic convergence to the Hannan-Quinn criterion with CHQ = 1.	Hannan-Quinn (1979)
8	$g_k = \delta\gamma^{(1/p_k)} / \left(1 - \delta\gamma^{(1/p_k)}\right)$	A natural conjugate prior structure, subjectively elicited through predictive implications. $\gamma < 1$ (so that g increases with k_j) and delta such that g/(1+g) Є [0.10, 0.15] (the weight of the "prior prediction error" in the Bayes factors); for k_j ranging from 1 to 15. FLS suggest covering this interval with the values of γ = 0.65 and δ = 0.15.	Laud and Ibrahim (1996)
9	$g = 1/p^2$	This prior is suggested by the risk inflation criterion (RIC).	Foster and George (1994)
10	$g = 1/(\max[n, p^2])$	The preferred prior by Fernandez Ley and Steel (2001), a mix of Prior 9 or Prior 1.	FLS (2001b)
11	$\beta \sim N(\mu, \sigma^2 V)$ $V = \sigma^2 \phi^2 (1/nX'X)^{-1}$ $v\lambda / \sigma^2 \sim \chi^2$	Data dependent priors. ϕ = 2.85, ν = 2.58, λ = 0.28 if the R^2 of the full model is less than 0.9, and ϕ = 9.2, ν = 0.2, λ = 0.1684 if the R^2 of the full model is greater than 0.9.	Raftery, Madigan and Hoeting (1997)
12	$g = n^{-1}$	Similar to the Unit Information Prior.	FLS(2001b)

Table 2
Posterior Inclusion Probabilities Across Parameter Priors
Model Prior = Uniform
(Growth Dataset)

	Priors Arranged By Effective g-Value (increasing left to right)										
	Prior 11	9 (FLS)	Prior 6	Prior 1 (UIP)	Prior 12	Prior 3	Prior 4	Prior 8	Prior 2	Prior 5	Prior 7
Confucius	99.5	99.9	100	100.0	100.0	100.0	100.0	100.0	99.9	99.2	98.5
GDPsh560	99.9	99.9	100	100.0	100.0	100.0	100.0	100.0	100.0	99.5	98.5
Life	96.5	96.4	99.9	100.0	100.0	99.9	99.8	98.6	96.4	93.1	90.9
RuleofLaw	47.2	64.0	99.6	100.0	99.6	99.6	98.3	93.0	69.3	57.3	56.6
SubSahara	74.8	83.8	99.9	100.0	100.0	100.0	99.7	97.5	86.3	80.2	79.6
EquipInv	99.0	96.8	98.3	99.9	98.4	98.3	95.6	88.8	94.4	95.3	95.2
Hindu	3.2	10.3	96.6	99.9	97.0	96.8	88.7	42.8	16.7	15.0	18.5
HighEnroll	0.3	0.7	93.4	99.8	94.0	93.5	78.1	2.8	2.1	3.9	7.2
LabForce	0.4	1.3	94.5	99.8	95.0	94.6	81.6	11.6	3.9	5.6	9.2
EthnoLFrac	0.5	1.3	90.8	99.3	91.4	90.8	74.6	7.2	3.3	4.8	8.0
Mining	28.0	38.5	96.4	99.2	96.5	96.4	93.3	74.7	49.1	43.4	44.1
LatAmerica	9.2	13.4	79.5	97.2	80.3	79.4	61.0	30.2	17.7	17.5	19.1
SpanishCol	0.0	0.1	67.6	94.6	68.7	67.3	42.3	2.0	0.5	1.1	2.4
FrenchCol	0.3	0.2	65.4	93.9	66.5	65.1	39.4	0.0	0.3	1.0	2.2
BritCol	0.0	0.0	64.7	93.6	65.8	64.4	38.7	0.7	0.2	0.6	1.8
PrSc	19.3	12.0	72.2	90.7	72.8	72.2	58.0	8.1	14.1	16.1	17.5
CivlLib	5.2	3.3	66.8	85.7	67.5	66.7	51.2	3.7	4.4	5.4	7.1
NEquipInv	28.8	49.3	71.3	85.6	71.7	71.3	66.6	82.1	52.4	41.1	40.3
English.	0.5	1.1	58	84.5	58.9	57.7	36.7	2.7	2.2	2.4	3.5
OutwarOr	0.0	0.0	51.2	82.8	52.2	51.0	31.4	0.7	0.2	0.6	1.7
BlMktPm	5.1	12.2	63.8	72.5	63.9	64.1	67.6	45.4	19.6	17.4	19.9
Muslim	66.9	68.3	44.3	60.9	44.4	44.4	49.4	54.9	66.5	60.3	56.1
Buddha	4.1	10.2	19.5	36.5	19.7	19.7	21.5	31.1	13.4	10.6	11.4
EcoOrg	34.2	56.6	39.5	35.6	39.2	39.7	50.1	88.7	61.0	47.3	45.2
X.PublEdu	0.0	0.2	17.9	13.3	17.8	18.1	19.4	1.5	0.6	1.1	2.0
PolRights	2.0	2.7	16.4	12.4	16.5	16.5	14.6	10.1	4.5	4.4	4.8
Protestants	35.5	51.5	25.7	11.7	25.2	26.0	41.7	81.3	56.8	47.7	46.4
WarDummy	1.1	0.9	6.2	11.7	6.4	6.3	3.9	0.8	1.2	1.8	2.0
Age	0.4	0.7	14.6	11.4	14.7	14.7	12.2	3.3	1.3	1.7	2.3
RFEXDist	1.8	2.0	4.6	9.6	4.7	4.7	4.0	0.6	2.6	3.3	3.4
Catholic	4.1	8.7	3.5	7.5	3.5	3.6	7.1	20.3	11.0	8.3	8.2
Popg	0.2	0.3	2.2	3.6	2.2	2.3	2.2	0.2	0.5	0.5	0.5
PrExports	2.2	2.5	1.2	2.8	1.2	1.2	2.1	5.9	3.7	3.0	2.8
Foreign.	0.5	0.3	0.7	2.0	0.7	0.7	0.4	0.0	0.2	0.6	0.7
Jewish	0.0	0.0	0.8	1.3	0.8	0.8	0.7	0.0	0.0	0.0	0.1
std.BMP.	0.0	0.0	0.6	1.3	0.6	0.6	0.4	0.0	0.0	0.0	0.0
Area	0.0	0.0	0.8	1.1	0.9	0.9	1.1	0.0	0.1	0.1	0.2
Work.Pop	0.4	0.2	0.3	1.1	0.3	0.3	0.2	0.0	0.2	0.6	0.8
AbsLat	0.6	0.5	1.2	1.0	1.2	1.2	1.8	0.3	0.7	0.9	1.0
YrsOpen	57.8	40.9	1.2	1.0	1.1	1.2	3.4	15.3	37.3	44.2	42.4
Rev.Coup	0.1	0.2	0.4	0.7	0.4	0.4	0.7	1.1	0.5	0.4	0.4
# of relevant regressors	7	8	15	22	21	21	17	10	10	7	7
Adj. R2 best models	0.827	0.843	0.845	0.925	0.915	0.915	0.874	0.846	0.825	0.745	0.685

1) Shaded cells indicated posterior inclusion probability over 50% (Jeffreys, 1961)
2) Priors 9 and 10 are identical in the growth context

Table 3
Best Models Across Parameter Priors
(Growth Dataset)

	Priors Arranged By Effective g-Value (increasing left to right)										
	Prior 11 Best Model	Prior 9 (FLS) Best Model	Prior 6 Best Model	Prior 1 (UIP) Best Model	Prior 12 Best Model	Prior 3 Best Model	Prior 4 Best Model	Prior 8 Best Model	Prior 2 Best Model	Prior 5 Best Model	Prior 7 Best Model
Confucius	0.0576	0.0575	0.0711	0.0759	0.071	0.0708	0.0708	0.0527	0.0499	0.0665	0.0381
GDPsh560	-0.0184	-0.0165	-0.0176	-0.0188	-0.0176	-0.0176	-0.0176	-0.0135	-0.0144	-0.0169	-0.0086
Life	0.0008	0.0008	0.0008	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0008	0.0007
RuleofLaw	0.0166	0.0168	0.0133	0.0131	0.0133	0.0133	0.0133	0.0122	0.0146	0.0107	.
SubSahara	-0.0166	-0.0133	-0.025	-0.0218	-0.025	-0.025	-0.025	-0.0117	-0.0115	-0.0154	.
EquipInv	0.154	0.1587	0.1511	0.1511	0.1509	0.1505	0.1505	0.0964	0.1378	0.1466	0.1441
Hindu	.	.	-0.1224	-0.1108	-0.1223	-0.122	-0.122	.	.	-0.0656	.
HighEnroll	.	.	-0.1313	-0.1213	-0.1312	-0.1308	-0.1308	.	.	-0.0894	.
LabForce	.	.	0.0000	0.0000	.	.	.	.	.	.	.
EthnoLFrac	.	.	0.0163	0.0165	0.0163	0.0163	0.0163	.	.	0.0143	.
Mining	.	.	0.0301	0.0328	0.03	0.03	0.03	0.028	.	0.0355	.
LatAmerica	.	.	-0.0159	-0.0127	-0.0159	-0.0158	-0.0158	.	.	.	.
SpanishCol	.	.	0.0161	0.0140	0.016	0.016	0.016	.	.	.	.
FrenchCol	.	.	0.0136	0.0110	0.0136	0.0136	0.0136	.	.	.	.
BritCol	.	.	0.01	0.0079	0.01	0.01	0.01	.	.	.	.
PrSc	.	.	0.0187	0.0249	0.0187	0.0187	0.0187	.	.	0.0253	.
CivlLib	.	.	-0.0026	-0.0028	-0.0026	-0.0026	-0.0026	.	.	-0.0021	.
NEquipInv	0.0603	0.0635	0.0377	0.0295	0.0376	0.0375	0.0375	0.0412	0.0552	.	.
English.	.	.	-0.0083	-0.0078	-0.0083	-0.0082	-0.0082	.	.	.	.
OutwarOr	.	.	-0.0038	-0.0035	-0.0038	-0.0038	-0.0038	.	.	.	.
BlMktPm	.	.	.	-0.0055	.	.	.	-0.0059	.	-0.0087	.
Muslim	0.0107	0.01	.	0.0078	.	.	.	0.0105	0.0087	0.0157	0.0078
Buddha	.	.	.	.	.	.	.	0.0103	.	.	.
EcoOrg	0.0029	0.0031	.	.	.	.	.	0.0024	0.0027	.	.
X.PublEdu	.	.	.	.	.	.	.	.	.	.	.
PolRights	.	.	.	.	.	.	.	.	.	.	.
Protestants	.	-0.011	.	.	.	.	.	-0.0082	-0.0096	.	-0.009
WarDummy	.	.	.	.	.	.	.	.	.	.	.
Age	.	.	.	.	.	.	.	.	.	.	.
RFEXDist	.	.	.	.	.	.	.	.	.	.	.
Catholic	.	.	.	.	.	.	.	.	.	.	.
Popg	.	.	.	.	.	.	.	.	.	.	.
PrExports	.	.	.	.	.	.	.	.	.	.	.
Foreign.	.	.	.	.	.	.	.	.	.	.	.
Jewish	.	.	.	.	.	.	.	.	.	.	.
std.BMP.	.	.	.	.	.	.	.	.	.	.	.
Area	.	.	.	.	.	.	.	.	.	.	.
Work.Pop	.	.	.	.	.	.	.	.	.	.	.
AbsLat	.	.	.	.	.	.	.	.	.	.	.
YrsOpen	.	.	.	.	.	.	.	.	.	.	0.0116
Rev.Coup	.	.	.	.	.	.	.	.	.	.	.
# of Regressors	**8**	**9**	**14**	**22**	**19**	**19**	**19**	**12**	**10**	**14**	**6**
R2	**0.846**	**0.863**	**0.876**	**0.948**	**0.938**	**0.938**	**0.938**	**0.872**	**0.85**	**0.899**	**0.712**
Adj. R2	**0.827**	**0.843**	**0.845**	**0.924**	**0.915**	**0.915**	**0.915**	**0.846**	**0.825**	**0.874**	**0.685**

1) Priors 9 and 10 are identical in the growth context

Table 4: Posterior Means, SDs, and Inclusion Probabilities Across Parameter Priors

Priors Arranged By Effective g-Value (increasing left to right)

	Prior 11			Prior 9 FLS			Prior 6			Prior 1 UIP			Prior 12			Prior 3		
	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD
Confucius	99.5	0.0553	0.0130	99.9	0.0559	0.0121	100.0	0.0688	0.0109	100.0	0.0732	0.0104	100.0	0.0688	0.0109	100.0	0.0685	0.0110
GDPsh560	99.9	-0.0163	0.0030	99.9	-0.0163	0.0029	100.0	-0.0174	0.0024	100.0	-0.0180	0.0023	100.0	-0.0174	0.0024	100.0	-0.0173	0.0024
Life	96.5	0.0008	0.0003	96.4	0.0008	0.0003	99.9	0.0009	0.0002	100.0	0.0009	0.0002	100.0	0.0009	0.0002	99.9	0.0009	0.0002
RuleofLaw	47.2	0.0073	0.0084	64.0	0.0100	0.0084	99.6	0.0123	0.0038	100.0	0.0124	0.0038	99.6	0.0123	0.0038	99.6	0.0123	0.0039
SubSahara	74.8	-0.0111	0.0076	83.8	-0.0129	0.0072	99.9	-0.0218	0.0048	100.0	-0.0222	0.0044	100.0	-0.0218	0.0047	100.0	-0.0217	0.0048
EquipInv	99.0	0.1871	0.0557	96.8	0.1644	0.0593	98.3	0.1403	0.0425	99.9	0.1466	0.0363	98.4	0.1403	0.0423	98.3	0.1396	0.0430
Hindu	3.2	-0.0009	0.0060	10.3	-0.0030	0.0104	96.6	-0.0940	0.0315	99.9	-0.1075	0.0205	97.0	-0.0947	0.0309	96.8	-0.0936	0.0314
HighEnroll	0.3	-0.0002	0.0051	0.7	-0.0007	0.0082	93.4	-0.1056	0.0408	99.8	-0.1193	0.0298	94.0	-0.1063	0.0402	93.5	-0.1052	0.0409
LabForce	0.4	0.0000	0.0000	1.3	0.0000	0.0000	94.5	0.0000	0.0000	99.8	0.0000	0.0000	95.0	0.0000	0.0000	94.6	0.0000	0.0000
EthnoLFrac	0.5	0.0001	0.0009	1.3	0.0001	0.0013	90.8	0.0131	0.0058	99.3	0.0152	0.0042	91.4	0.0132	0.0057	90.8	0.0130	0.0058
Mining	28.0	0.0113	0.0198	38.5	0.0151	0.0212	96.4	0.0334	0.0134	99.2	0.0331	0.0117	96.5	0.0333	0.0134	96.4	0.0333	0.0136
LatAmerica	9.2	-0.0008	0.0027	13.4	-0.0012	0.0033	79.5	-0.0105	0.0066	97.2	-0.0129	0.0048	80.3	-0.0107	0.0066	79.4	-0.0105	0.0067
SpanishCol	0.0	0.0000	0.0001	0.1	0.0000	0.0003	67.6	0.0098	0.0076	94.6	0.0134	0.0051	68.7	0.0099	0.0075	67.3	0.0097	0.0076
FrenchCol	0.3	0.0000	0.0004	0.2	0.0000	0.0004	65.4	0.0076	0.0061	93.9	0.0106	0.0041	66.5	0.0078	0.0061	65.1	0.0076	0.0061
BritCol	0.0	0.0000	0.0000	0.0	0.0000	0.0001	64.7	0.0052	0.0043	93.6	0.0073	0.0032	65.8	0.0052	0.0043	64.4	0.0051	0.0043
PrSc	19.3	0.0043	0.0096	12.0	0.0026	0.0077	72.2	0.0157	0.0116	90.7	0.0194	0.0093	72.8	0.0158	0.0115	72.2	0.0156	0.0116
CivlLib	5.2	-0.0001	0.0006	3.3	-0.0001	0.0005	66.8	-0.0017	0.0014	85.7	-0.0023	0.0012	67.5	-0.0017	0.0014	66.7	-0.0017	0.0014
NEquipInv	28.8	0.0169	0.0286	49.3	0.0292	0.0327	71.3	0.0303	0.0247	85.6	0.0320	0.0210	71.7	0.0303	0.0245	71.3	0.0302	0.0247
English.	0.5	0.0000	0.0005	1.1	-0.0001	0.0008	58.0	-0.0048	0.0049	84.5	-0.0067	0.0045	58.9	-0.0048	0.0049	57.7	-0.0047	0.0049
OutwarOr	0.0	0.0000	0.0000	0.0	0.0000	0.0001	51.2	-0.0019	0.0022	82.8	-0.0031	0.0021	52.2	-0.0019	0.0022	51.0	-0.0019	0.0022
BlMktPm	5.1	-0.0004	0.0020	12.2	-0.0009	0.0028	63.8	-0.0046	0.0043	72.5	-0.0043	0.0036	63.9	-0.0046	0.0043	64.1	-0.0047	0.0044
Muslim	66.9	0.0086	0.0071	68.3	0.0083	0.0066	44.3	0.0051	0.0067	60.9	0.0054	0.0056	44.4	0.0050	0.0067	44.4	0.0051	0.0067
Buddha	4.1	0.0005	0.0027	10.2	0.0013	0.0041	19.5	0.0015	0.0037	36.5	0.0027	0.0044	19.7	0.0015	0.0037	19.7	0.0015	0.0037
EcoOrg	34.2	0.0010	0.0014	56.6	0.0016	0.0015	39.5	0.0007	0.0011	35.6	0.0005	0.0008	39.2	0.0007	0.0011	39.7	0.0007	0.0011
X.PublEdu	0.0	0.0000	0.0034	0.2	0.0004	0.0098	17.9	0.0395	0.0963	13.3	0.0212	0.0688	17.8	0.0392	0.0960	18.1	0.0399	0.0968
PolRights	2.0	0.0000	0.0003	2.7	0.0000	0.0003	16.4	-0.0003	0.0007	12.4	-0.0002	0.0007	16.5	-0.0003	0.0007	16.5	-0.0003	0.0007
Protestants	35.5	-0.0045	0.0066	51.5	-0.0065	0.0071	25.7	-0.0028	0.0053	11.7	-0.0009	0.0031	25.2	-0.0027	0.0053	26.0	-0.0028	0.0053
WarDummy	1.1	-0.0001	0.0005	0.9	0.0000	0.0005	6.2	-0.0002	0.0008	11.7	-0.0003	0.0010	6.4	-0.0002	0.0008	6.3	-0.0002	0.0008
Age	0.4	0.0000	0.0000	0.7	0.0000	0.0000	14.6	0.0000	0.0000	11.4	0.0000	0.0000	14.7	0.0000	0.0000	14.7	0.0000	0.0000
RFEXDist	1.8	0.0000	0.0000	2.0	0.0000	0.0000	4.6	0.0000	0.0000	9.6	0.0000	0.0000	4.7	0.0000	0.0000	4.7	0.0000	0.0000
Catholic	4.1	-0.0003	0.0017	8.7	-0.0006	0.0025	3.5	-0.0001	0.0013	7.5	0.0003	0.0014	3.5	-0.0001	0.0013	3.6	-0.0001	0.0013
Popg	0.2	0.0005	0.0128	0.3	0.0006	0.0141	2.2	0.0013	0.0318	3.6	-0.0034	0.0366	2.2	0.0012	0.0321	2.3	0.0014	0.0324
PrExports	2.2	-0.0003	0.0019	2.5	-0.0003	0.0019	1.2	-0.0001	0.0008	2.8	-0.0001	0.0010	1.2	-0.0001	0.0007	1.2	-0.0001	0.0008
Foreign.	0.5	0.0000	0.0005	0.3	0.0000	0.0004	0.7	0.0000	0.0003	2.0	0.0000	0.0004	0.7	0.0000	0.0003	0.7	0.0000	0.0003
Jewish	0.0	0.0000	0.0000	0.0	0.0000	0.0001	0.8	0.0000	0.0008	1.3	0.0000	0.0010	0.8	0.0000	0.0008	0.8	0.0000	0.0008
std.BMP.	0.0	0.0000	0.0000	0.0	0.0000	0.0000	0.6	0.0000	0.0000	1.3	0.0000	0.0000	0.6	0.0000	0.0000	0.6	0.0000	0.0000
Area	0.0	0.0000	0.0000	0.0	0.0000	0.0000	0.8	0.0000	0.0000	1.1	0.0000	0.0000	0.9	0.0000	0.0000	0.9	0.0000	0.0000
Work.Pop	0.4	-0.0001	0.0009	0.2	0.0000	0.0007	0.3	0.0000	0.0003	1.1	0.0000	0.0007	0.3	0.0000	0.0003	0.3	0.0000	0.0003
AbsLat	0.6	0.0000	0.0000	0.5	0.0000	0.0000	1.2	0.0000	0.0000	1.0	0.0000	0.0000	1.2	0.0000	0.0000	1.2	0.0000	0.0000
YrsOpen	57.8	0.0088	0.0083	40.9	0.0061	0.0080	1.2	0.0001	0.0010	1.0	0.0000	0.0004	1.1	0.0001	0.0009	1.2	0.0001	0.0009
Rev.Coup	0.1	0.0000	0.0002	0.2	0.0000	0.0003	0.4	0.0000	0.0002	0.7	0.0000	0.0003	0.4	0.0000	0.0002	0.4	0.0000	0.0002

	Prior 4			Prior 8			Prior 2			Prior 5			Prior 7		
	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD	Prob Coef ≠ 0	Posterior Mean	Posterior SD
Confucius	100.0	0.0600	0.0142	100.0	0.0483	0.0140	99.9	0.0488	0.0141	99.2	0.0408	0.0158	98.5	0.0361	0.0164
GDPsh560	100.0	-0.0155	0.0030	100.0	-0.0141	0.0030	100.0	-0.0142	0.0031	99.5	-0.0116	0.0034	98.5	-0.0101	0.0036
Life	99.8	0.0008	0.0003	98.6	0.0007	0.0003	96.4	0.0007	0.0003	93.1	0.0006	0.0003	90.9	0.0005	0.0003
RuleofLaw	98.3	0.0112	0.0052	93.0	0.0119	0.0062	69.3	0.0092	0.0077	57.3	0.0063	0.0071	56.6	0.0054	0.0066
SubSahara	99.7	-0.0183	0.0056	97.5	-0.0143	0.0058	86.3	-0.0118	0.0067	80.2	-0.0093	0.0066	79.6	-0.0083	0.0063
EquipInv	95.6	0.1203	0.0555	88.8	0.0978	0.0616	94.4	0.1335	0.0646	95.3	0.1234	0.0671	95.2	0.1107	0.0669
Hindu	88.7	-0.0679	0.0387	42.8	-0.0116	0.0185	16.7	-0.0046	0.0133	15.0	-0.0042	0.0138	18.5	-0.0055	0.0163
HighEnroll	78.1	-0.0758	0.0528	2.8	-0.0022	0.0145	2.1	-0.0016	0.0124	3.9	-0.0025	0.0153	7.2	-0.0043	0.0197
LabForce	81.6	0.0000	0.0000	11.6	0.0000	0.0000	3.9	0.0000	0.0000	5.6	0.0000	0.0000	9.2	0.0000	0.0000
EthnoLFrac	74.6	0.0095	0.0072	7.2	0.0006	0.0026	3.3	0.0003	0.0020	4.8	0.0004	0.0022	8.0	0.0006	0.0028
Mining	93.3	0.0309	0.0173	74.7	0.0236	0.0203	49.1	0.0166	0.0209	43.4	0.0123	0.0189	44.1	0.0109	0.0178
LatAmerica	61.0	-0.0069	0.0069	30.2	-0.0024	0.0042	17.7	-0.0014	0.0034	17.5	-0.0012	0.0031	19.1	-0.0012	0.0030
SpanishCol	42.3	0.0054	0.0073	2.0	0.0001	0.0010	0.5	0.0000	0.0006	1.1	0.0001	0.0010	2.4	0.0002	0.0015
FrenchCol	39.4	0.0041	0.0057	0.0	0.0000	0.0001	0.3	0.0000	0.0004	1.0	0.0001	0.0008	2.2	0.0001	0.0012
BritCol	38.7	0.0027	0.0040	0.7	0.0000	0.0003	0.2	0.0000	0.0002	0.6	0.0000	0.0005	1.8	0.0001	0.0008
PrSc	58.0	0.0119	0.0122	8.1	0.0014	0.0056	14.1	0.0027	0.0075	16.1	0.0027	0.0074	17.5	0.0026	0.0071
CivlLib	51.2	-0.0011	0.0014	3.7	-0.0001	0.0004	4.5	-0.0001	0.0005	5.4	-0.0001	0.0005	7.1	-0.0001	0.0006
NEquipInv	66.6	0.0283	0.0278	82.1	0.0402	0.0288	52.4	0.0264	0.0308	41.1	0.0170	0.0267	40.3	0.0144	0.0248
English.	36.7	-0.0027	0.0046	2.7	-0.0001	0.0011	2.2	-0.0001	0.0011	2.4	-0.0001	0.0011	3.5	-0.0001	0.0014
OutwarOr	31.4	-0.0010	0.0020	0.7	0.0000	0.0002	0.2	0.0000	0.0002	0.6	0.0000	0.0003	1.7	0.0000	0.0005
BlMktPm	67.6	-0.0049	0.0046	45.4	-0.0028	0.0041	19.6	-0.0013	0.0032	17.4	-0.0010	0.0028	19.9	-0.0010	0.0028
Muslim	49.4	0.0059	0.0072	54.9	0.0058	0.0062	66.5	0.0072	0.0064	60.3	0.0056	0.0060	56.1	0.0046	0.0056
Buddha	21.5	0.0018	0.0044	31.1	0.0031	0.0059	13.4	0.0014	0.0043	10.6	0.0009	0.0036	11.4	0.0009	0.0036
EcoOrg	50.1	0.0010	0.0012	88.7	0.0021	0.0012	61.0	0.0015	0.0014	47.3	0.0009	0.0012	45.2	0.0008	0.0011
X.PublEdu	19.4	0.0396	0.0995	1.5	0.0023	0.0245	0.6	0.0011	0.0175	1.1	0.0016	0.0223	2.0	0.0026	0.0291
PolRights	14.6	-0.0002	0.0007	10.1	-0.0001	0.0005	4.5	-0.0001	0.0004	4.4	-0.0001	0.0003	4.8	0.0000	0.0003
Protestants	41.7	-0.0043	0.0062	81.3	-0.0087	0.0065	56.8	-0.0062	0.0068	47.7	-0.0043	0.0061	46.4	-0.0038	0.0057
WarDummy	3.9	-0.0001	0.0007	0.8	0.0000	0.0004	1.2	0.0000	0.0005	1.8	-0.0001	0.0006	2.0	-0.0001	0.0006
Age	12.2	0.0000	0.0000	3.3	0.0000	0.0000	1.3	0.0000	0.0000	1.7	0.0000	0.0000	2.3	0.0000	0.0000
RFEXDist	4.0	0.0000	0.0000	0.6	0.0000	0.0000	2.6	0.0000	0.0000	3.3	0.0000	0.0000	3.4	0.0000	0.0000
Catholic	7.1	-0.0003	0.0020	20.3	-0.0012	0.0034	11.0	-0.0007	0.0026	8.3	-0.0004	0.0020	8.2	-0.0003	0.0018
Popg	2.2	0.0023	0.0365	0.2	0.0004	0.0125	0.5	0.0010	0.0188	0.5	0.0007	0.0175	0.5	0.0006	0.0181
PrExports	2.1	-0.0001	0.0012	5.9	-0.0004	0.0023	3.7	-0.0003	0.0021	3.0	-0.0002	0.0017	2.8	-0.0002	0.0016
Foreign.	0.4	0.0000	0.0002	0.0	0.0000	0.0001	0.2	0.0000	0.0003	0.6	0.0000	0.0005	0.7	0.0000	0.0004
Jewish	0.7	0.0000	0.0010	0.0	0.0000	0.0000	0.0	0.0000	0.0002	0.0	0.0000	0.0003	0.1	0.0000	0.0004
std.BMP.	0.4	0.0000	0.0000	0.0	0.0000	0.0000	0.0	0.0000	0.0000	0.0	0.0000	0.0000	0.0	0.0000	0.0000
Area	1.1	0.0000	0.0000	0.0	0.0000	0.0000	0.1	0.0000	0.0000	0.1	0.0000	0.0000	0.2	0.0000	0.0000
Work.Pop	0.2	0.0000	0.0004	0.0	0.0000	0.0001	0.2	0.0000	0.0006	0.6	-0.0001	0.0010	0.8	-0.0001	0.0011
AbsLat	1.8	0.0000	0.0000	0.3	0.0000	0.0000	0.7	0.0000	0.0000	0.9	0.0000	0.0000	1.0	0.0000	0.0000
YrsOpen	3.4	0.0003	0.0018	15.3	0.0015	0.0042	37.3	0.0047	0.0070	44.2	0.0049	0.0067	42.4	0.0042	0.0062
Rev.Coup	0.7	0.0000	0.0004	1.1	0.0000	0.0006	0.5	0.0000	0.0004	0.4	0.0000	0.0004	0.4	0.0000	0.0004

1) Priors 9 and 10 are identical in the growth context

Table 5
Sampler Comparsion:
Best Model, Posterior Inclusion Probabilities and System Time

Sampler	**Branch and Bounds**		**MC3 (FLS)**[b]		**Coinflip**[a]		**PcGets**[c]
Prior/g value	***prior 1***		***g=1/n***		***prior 1***		***NA***
Computer Language	**R**		**Fortran77**		**GAUSS**		**GiveWin**
	Posterior	**Best model**	**Posterior**	**Best model**	**Posterior**	**Best model**	**PcGets**
Confucius	100.0	0.0759	100.0	0.0720	100.0	0.0571	0.056
GDPsh560	100.0	-0.0188	100.0	-0.0179	100.0	-0.0164	-0.165
Life	100.0	0.0009	99.9	0.0009	99.9	0.0010	0.098
RuleofLaw	100.0	0.0131	96.3	0.0135	92.1	0.0119	0.015
SubSahara	100.0	-0.0218	99.8	-0.0254	99.7	-0.0228	-0.027
EquipInv	99.9	0.1511	98.0	0.1530	98.1		0.186
Hindu	99.9	-0.1108	98.5	-0.1240	99.4	-0.0759	-0.111
HighEnroll	99.8	-0.1213	95.2	-0.1330	97.6	-0.0701	-0.114
LabForce	99.8	0.0000	96.8	0.0000	99.0	0.0000	0.004
EthnoLFrac	99.3	0.0165	91.1	0.0166	95.3		0.013
Mining	99.2	0.0328	95.2	0.0305	97.1	0.0431	0.034
LatAmerica	97.2	-0.0127	78.9	-0.0161	89.1	-0.0090	-0.016
SpanishCol	94.6	0.0140	63.8	0.0163	81.8		0.015
FrenchCol	93.9	0.0110	58.8	0.0138	78.5		0.011
BritCol	93.6	0.0079	53.3	0.0101	71.2		0.008
PrSc	90.7	0.0249	64.2	0.0190	77.6		0.017
CivlLib	85.7	-0.0028	55.1	-0.0027	67.7		
NEquipInv	85.6	0.0295	72.7	0.0382	82.0	0.0560	
English.	84.5	-0.0078	45.1	-0.0084	60.2		
OutwarOr	82.8	-0.0035	43.1	-0.0039	68.1		
BlMktPm	72.5	-0.0055	67.8		73.2	-0.0090	
Muslim	60.9	0.0078	40.7		52.6		
Buddha	36.5		21.2		39.7		
EcoOrg	35.6		54.5		56.5	0.0023	
X.PublEdu	13.3		28.1		37.0	0.2671	
PolRights	12.4		28.9		35.9	-0.0015	
Protestants	11.7		41.7		45.5	-0.0152	
WarDummy	11.7		9.6		29.7	-0.0028	
Age	11.4		23.6		33.0	0.0000	
RFEXDist	9.6		3.1		23.8		
Catholic	7.5		7.1		21.0	-0.0033	
Popg	3.6		5.5		17.0	0.2119	
PrExports	2.8		3.1		16.0		
Foreign.	2.0		3.0		15.0		
Jewish	1.3		2.5		13.4		
std.BMP.	1.3		2.9		13.8		
Area	1.1		3.8		14.1		
Work.Pop	1.1		3.0		13.5		
AbsLat	1.0		4.7		15.0		
YrsOpen	1.0		4.1		15.4		
Rev.Coup	0.7		8.4		13.1		
# of Regressors		**22**		**20**		**20**	**17**
BIC		**-118.161**		**-117.490**		**-100.356**	**-118.014**
R2		**0.948**		**0.940**		**0.924**	**0.929**
Adj. R2		**0.924**		**0.917**		**0.895**	**0.907**
Best model is BMA.COMPARE's best model #		**1**		**3**		**not in top 5**	**not in top 5**
System time in seconds (h/min)	**798sec (0.22h)**		**1555sec (0.43h)**		**15359sec (4.25h)**		**900sec (0.25h)**[c]

[a]BACE defaults used, see http://www.econ.cam.ac.uk/faculty/doppelhofer/. These defaults assured that the program does converged with this growth dataset

[b]FLS Fortran77 defaults used, see http://qed.econ.queensu.ca/jae/2001-v16.5/fernandez-ley-steel/. The result is robust to quadrupling the default number of integer chains (the maximum for the test computer)

[c] PcGets (Hendry and Krolzig, 2004) reports only the best regression

Benchmarks for Dell OptiPlex GX270, Pentium 4, 3 GHz, 1 GB RAM, prior 1

Table 6

Posterior Inclusion Probabilities Across Parameter and Model Priors

Uniform Model Prior Column 1, All Other Columns: Prior Model Size =7 (as in Sala-i-Martin et al., 2004)

(Growth Dataset)

		Priors Arranged By Effective g-Value (increasing left to right)										
	Prior 1 UIP Model Prior: Uniform	Prior 11	Prior 9	Prior 6	Prior 1	Prior 12	Prior 3	Prior 4	Prior 8	Prior 2	Prior 5	Prior 7
Confucius	100.0	0.1	95.8	99.7	99.9	99.7	99.7	98.7	97.2	96.5	87.1	84.8
GDPsh560	100.0	0.0	91.7	99.8	100.0	99.8	99.8	99.0	97.3	96.8	71.8	50.1
Life	100.0	0.0	77.4	94.8	97.8	94.9	94.8	90.2	84.9	82.0	48.8	30.8
RuleofLaw	100.0	0.0	16.9	49.4	68.6	50.2	50.4	37.0	29.2	21.5	12.3	8.2
SubSahara	100.0	0.0	60.4	76.5	86.3	76.9	77.0	70.1	66.1	62.9	48.5	35.1
EquipInv	99.9	0.2	99.4	98.2	99.2	98.1	98.0	98.5	98.7	99.0	98.5	97.9
Hindu	99.9	0.0	0.0	4.8	9.6	5.0	5.1	2.3	1.1	0.1	0.0	0.0
HighEnroll	99.8	0.0	0.1	0.1	1.0	0.1	0.1	0.1	0.1	0.1	0.8	1.2
LabForce	99.8	0.0	0.0	0.3	1.5	0.3	0.3	0.1	0.0	0.0	0.0	0.0
EthnoLFrac	99.3	0.0	0.2	0.4	0.9	0.5	0.5	0.4	0.4	0.4	0.5	0.3
Mining	99.2	0.0	6.9	31.2	33.7	31.8	32.2	25.8	19.6	12.0	3.8	1.7
LatAmerica	97.2	0.0	6.0	11.2	11.1	11.4	11.6	11.6	10.9	9.3	6.1	3.9
SpanishCol	94.6	0.0	0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
FrenchCol	93.9	0.0	0.3	0.3	0.0	0.3	0.3	0.6	0.7	0.7	0.3	0.1
BritCol	93.6	0.0	0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PrSc	90.7	0.0	7.8	13.3	8.0	13.3	13.5	14.6	13.6	11.5	6.5	4.8
CivlLib	85.7	0.0	1.2	3.2	2.2	3.2	3.3	3.3	2.9	2.1	0.6	0.4
NEquipInv	85.6	0.0	5.6	34.7	56.2	35.4	35.5	23.0	16.6	9.8	5.2	4.1
English.	84.5	0.0	0.0	0.8	0.1	0.8	0.9	0.7	0.4	0.1	0.1	0.3
OutwarOr	82.8	0.0	0.0	0	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.3
BlMktPm	72.5	0.0	0.3	6.8	10.0	7.1	7.3	4.6	2.7	0.8	0.1	0.0
Muslim	60.9	0.0	29.2	65.6	69.1	65.9	65.8	56.5	46.9	37.2	13.0	7.2
Buddha	36.5	0.0	2.6	5.9	11.8	6.1	6.2	3.8	3.1	2.0	9.6	13.8
EcoOrg	35.6	0.0	7.6	40.7	61.9	41.6	41.7	27.4	19.7	11.9	6.2	5.0
X.PublEdu	13.3	0.0	0.0	0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PolRights	12.4	0.0	0.5	1.9	0.8	1.9	2.0	2.0	1.7	1.2	0.4	0.5
Protestants	11.7	0.0	21.3	40.7	51.8	41.3	41.5	32.6	27.4	21.4	24.9	25.6
WarDummy	11.7	0.0	0.9	1.2	0.0	1.2	1.2	1.9	2.1	1.9	1.3	0.7
Age	11.4	0.0	0.6	0.6	0.1	0.6	0.7	0.9	1.1	1.0	1.8	2.0
RFEXDist	9.6	0.0	1.6	2.5	0.0	2.5	2.6	3.3	3.3	2.6	3.8	4.8
Catholic	7.5	0.0	1.1	5.3	9.0	5.5	5.5	3.3	2.3	1.4	1.9	1.6
Popg	3.6	0.0	0.0	0.2	0.0	0.2	0.3	0.2	0.1	0.0	0.1	0.2
PrExports	2.8	0.0	0.1	1.8	1.3	1.9	1.9	1.4	0.9	0.3	0.5	0.5
Foreign.	2.0	0.0	0.9	0.6	0.0	0.6	0.6	1.1	1.3	1.5	1.0	0.7
Jewish	1.3	0.0	0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2
std.BMP.	1.3	0.0	0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.8
Area	1.1	0.0	0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3
Work.Pop	1.1	0.0	1.2	0.5	0.1	0.4	0.5	1.0	1.5	1.7	2.2	2.2
AbsLat	1.0	0.0	0.3	0.6	0.0	0.6	0.6	0.8	0.8	0.7	0.2	1.0
YrsOpen	1.0	0.0	63.0	52.4	38.0	51.8	51.7	59.2	61.0	63.5	49.1	38.2
Rev.Coup	0.7	0.0	0.0	0.1	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0
effect threshold 50%	22	6	6	7	10	8	8	7	6	6	3	3
effect threshold 17.08%	NA	7	8	12	12	12	12	12	11	9	7	7

1) Light shaded cells are inclusion probabilities > 50%. Dark shaded cells indicate the additional regressors that pass the Sala-i-Martin et al 2004 17.08% effect-threshold.

2) Priors 9 and 10 are identical in the growth context

Table 7

Parameter Priors And Predictive Performance

Performance Scores *Relative* to Parameter Prior 1

(Growth Dataset, Model Prior: Uniform)

Trials: 578

Prior	Median[a]	+ / 100[b]	Significance[c]
		MSE	
11	0.014	69	0.00
9	0.012	69	0.00
6	0.006	70	0.00
12	0.006	71	0.00
3	0.006	70	0.00
4	0.003	57	0.00
8	0.002	57	0.00
2	0.003	57	0.00
5	0.002	53	0.07
7	0.002	55	0.01
		CRPS	
11	0.030	68	0.00
9	0.032	68	0.00
6	0.008	64	0.00
12	0.009	65	0.00
3	0.008	64	0.00
4	0.002	53	0.07
8	0.003	55	0.01
2	0.007	57	0.00
5	0.011	59	0.00
7	0.021	64	0.00
		LPS	
11	0.969	62	0.00
9	1.540	65	0.00
6	1.478	75	0.00
12	1.748	78	0.00
3	1.478	75	0.00
4	-0.833	37	1.00
8	-0.861	38	1.00
2	-0.468	43	1.00
5	0.003	50	0.52
7	0.580	55	0.01

[a] Median refers to the median inprovement in the score attained by the UIP compared to a given alternative prior

[b] Indicates number of successes per 100 trials where "success" is a better predictive score by the UIP than by the alternative prior

[c] Significance refers to the binomial p values, P(X > or = z), for the given number of trials and successes; where success is defined as a better score for prior 1 as compared to a given alternative prior

1) Priors 9 and 10 are identical in the simulated dataset

2) Priors arranged by effective g-value (increasing top to bottom)

Table 8

Parameter Priors, Model Priors, and Predictive Performance (Growth Dataset)

Performance Scores *Relative* to Prior 1 with Uniform Model Prior

Trials = 190

Prior	Prior Model Size=3			Prior Model Size=5			Prior Model Size=6			Prior Model Size=7			Prior Model Size=8		
	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]
	MSE			**MSE**			**MSE**			**MSE**			**MSE**		
11	0.16	71	0.000	0.13	72	0.00	0.14	77	0.00	0.13	71	0.00	0.12	71	0.00
9	0.15	71	0.000	0.12	71	0.00	0.14	78	0.00	0.12	71	0.00	0.12	71	0.00
6	0.09	70	0.000	0.08	71	0.00	0.13	75	0.00	0.09	73	0.00	0.09	72	0.00
1	0.03	67	0.000	0.02	67	0.00	0.02	67	0.00	0.02	69	0.00	0.02	68	0.00
12	0.10	70	0.000	0.09	71	0.00	0.13	75	0.00	0.09	72	0.00	0.09	73	0.00
3	0.09	70	0.000	0.08	71	0.00	0.13	75	0.00	0.09	73	0.00	0.09	72	0.00
4	0.08	68	0.000	0.06	67	0.00	0.09	70	0.00	0.05	65	0.00	0.05	65	0.00
8	0.09	67	0.000	0.08	65	0.00	0.09	68	0.00	0.05	65	0.00	0.05	65	0.00
2	0.09	67	0.000	0.07	68	0.00	0.10	70	0.00	0.07	68	0.00	0.06	67	0.00
5	0.15	70	0.000	0.13	69	0.00	0.15	76	0.00	0.11	67	0.00	0.11	67	0.00
7	0.19	75	0.000	0.17	73	0.00	0.18	80	0.00	0.16	73	0.00	0.15	72	0.00
	CRPS			**CRPS**			**CRPS**			**CRPS**			**CRPS**		
11	0.04	83	0.000	0.04	80	0.00	0.03	78	0.00	0.03	79	0.00	0.02	77	0.00
9	0.04	85	0.000	0.03	77	0.00	0.02	79	0.00	0.02	79	0.00	0.02	75	0.00
6	0.01	69	0.000	0.01	66	0.00	0.01	69	0.00	0.01	66	0.00	0.01	67	0.00
1	0.00	61	0.002	0.00	59	0.01	0.00	62	0.00	0.00	61	0.00	0.00	61	0.00
12	0.02	73	0.000	0.01	68	0.00	0.01	71	0.00	0.01	71	0.00	0.01	69	0.00
3	0.01	69	0.000	0.01	66	0.00	0.01	69	0.00	0.01	66	0.00	0.01	67	0.00
4	0.01	67	0.000	0.00	56	0.05	0.00	59	0.01	0.00	56	0.06	0.00	53	0.21
8	0.01	65	0.000	0.00	57	0.02	0.00	59	0.01	0.00	56	0.05	0.00	56	0.06
2	0.01	72	0.000	0.01	66	0.00	0.01	66	0.00	0.00	58	0.01	0.00	57	0.03
5	0.01	65	0.000	0.01	65	0.00	0.00	63	0.00	0.00	61	0.00	0.00	61	0.00
7	0.01	63	0.000	0.01	58	0.01	0.00	60	0.00	0.00	59	0.01	0.00	62	0.00
	LPS			**LPS**			**LPS**			**LPS**			**LPS**		
11	1.50	62	0.00	1.17	59	0.01	1.89	64	0.00	1.23	58	0.01	1.15	57	0.02
9	1.53	62	0.00	1.31	60	0.00	1.96	64	0.00	1.30	59	0.01	1.16	59	0.01
6	0.40	54	0.14	0.66	54	0.17	1.72	59	0.01	0.68	55	0.11	0.89	56	0.05
1	0.86	62	0.00	0.57	61	0.00	0.43	61	0.00	0.34	60	0.00	0.42	59	0.01
12	0.66	56	0.06	0.76	54	0.17	1.99	59	0.01	0.87	57	0.03	1.13	57	0.03
3	0.40	54	0.14	0.66	54	0.17	1.72	59	0.01	0.68	55	0.11	0.89	56	0.05
4	0.18	52	0.31	-0.28	47	0.79	0.92	57	0.03	-0.44	47	0.83	-0.44	47	0.79
8	0.45	53	0.26	-0.05	48	0.69	0.58	57	0.03	-0.43	47	0.79	-0.40	47	0.83
2	0.46	53	0.26	0.10	51	0.41	0.99	58	0.01	-0.17	48	0.74	-0.28	48	0.69
5	1.45	59	0.01	1.11	58	0.01	1.68	61	0.00	0.81	56	0.05	0.69	55	0.08
7	1.86	62	0.00	1.69	61	0.00	2.06	64	0.00	1.45	58	0.01	1.30	58	0.01

Prior	Prior Model Size=9			Prior Model Size=11			Prior Model Size=13			Prior Model Size=15			Prior Model Size=17		
	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]	Median[a]	+ / 100[b]	Sig[c]
	MSE			**MSE**			**MSE**			**MSE**			**MSE**		
11	0.16	79	0.00	0.17	81	0.00	0.11	71	0.00	0.10	71	0.00	0.10	72	0.00
9	0.16	81	0.00	0.16	80	0.00	0.11	73	0.00	0.11	74	0.00	0.10	73	0.00
6	0.15	79	0.00	0.12	77	0.00	0.08	80	0.00	0.08	81	0.00	0.07	79	0.00
1	0.04	58	0.01	0.02	69	0.00	0.01	68	0.00	0.01	69	0.00	0.01	70	0.00
12	0.15	80	0.00	0.12	78	0.00	0.08	82	0.00	0.08	81	0.00	0.07	81	0.00
3	0.15	79	0.00	0.12	77	0.00	0.08	80	0.00	0.08	81	0.00	0.07	79	0.00
4	0.13	78	0.00	0.11	76	0.00	0.06	69	0.00	0.06	73	0.00	0.05	73	0.00
8	0.13	74	0.00	0.12	76	0.00	0.05	66	0.00	0.05	71	0.00	0.05	71	0.00
2	0.12	74	0.00	0.14	76	0.00	0.05	63	0.00	0.06	65	0.00	0.05	65	0.00
5	0.17	73	0.00	0.17	77	0.00	0.09	65	0.00	0.09	64	0.00	0.08	65	0.00
7	0.20	75	0.00	0.20	80	0.00	0.13	69	0.00	0.12	69	0.00	0.10	68	0.00
	CRPS			**CRPS**			**CRPS**			**CRPS**			**CRPS**		
11	0.03	72	0.00	0.05	78	0.00	0.01	77	0.00	0.01	72	0.00	0.01	71	0.00
9	0.02	71	0.00	0.04	75	0.00	0.01	73	0.00	0.01	74	0.00	0.01	71	0.00
6	0.01	62	0.00	0.01	63	0.00	0.00	71	0.00	0.00	62	0.00	0.00	63	0.00
1	0.01	53	0.26	0.00	59	0.01	0.00	57	0.03	0.00	57	0.03	0.00	55	0.08
12	0.01	62	0.00	0.01	63	0.00	0.00	68	0.00	0.00	65	0.00	0.00	68	0.00
3	0.01	62	0.00	0.01	63	0.00	0.00	71	0.00	0.00	62	0.00	0.00	63	0.00
4	0.01	59	0.01	0.00	53	0.15	0.00	59	0.01	0.00	57	0.03	0.00	57	0.02
8	0.00	55	0.11	0.00	56	0.01	0.00	53	0.26	0.00	55	0.11	0.00	56	0.06
2	0.01	61	0.00	0.01	65	0.00	0.00	53	0.21	0.00	54	0.17	0.00	52	0.31
5	0.01	59	0.01	0.01	57	0.00	0.00	55	0.11	0.00	51	0.47	0.00	51	0.41
7	0.01	60	0.00	0.01	57	0.00	0.00	55	0.08	0.00	52	0.36	0.00	51	0.47
	LPS			**LPS**			**LPS**			**LPS**			**LPS**		
11	3.58	82	0.00	4.18	78	0.00	0.82	55	0.08	0.89	57	0.03	0.99	57	0.02
9	3.51	82	0.00	4.28	78	0.00	1.10	56	0.06	1.21	59	0.01	1.24	61	0.00
6	2.79	82	0.00	3.23	75	0.00	1.26	65	0.00	1.41	70	0.00	1.60	72	0.00
1	1.63	60	0.00	0.32	61	0.00	0.22	61	0.00	0.18	62	0.00	0.10	63	0.00
12	2.89	83	0.00	3.34	76	0.00	1.46	67	0.00	1.58	72	0.00	1.94	73	0.00
3	2.79	82	0.00	3.23	75	0.00	1.26	65	0.00	1.41	70	0.00	1.60	72	0.00
4	2.36	77	0.00	3.01	72	0.00	-0.54	44	0.97	-0.48	45	0.94	-0.53	43	0.98
8	2.37	77	0.00	3.11	73	0.00	-0.62	42	0.99	-0.61	43	0.98	-0.82	44	0.95
2	2.65	78	0.00	3.29	75	0.00	-0.39	48	0.74	-0.40	46	0.89	-0.54	45	0.94
5	3.43	88	0.00	4.05	79	0.00	0.20	52	0.31	0.02	50	0.53	-0.02	50	0.53
7	4.22	91	0.00	4.61	81	0.00	0.83	55	0.08	0.69	54	0.14	0.57	53	0.21

[a] Median refers to the median inprovement in the score attained by the UIP compared to a given alternative

[b] Indicates number of successes per 100 trials where "success" is a better predictive score by the UIP than by the alternative prior

[c] Significance refers to the binomial p values P(X > or = z) for the given number of trials and successes;

1) Priors 9 and 10 are identical in the simulated dataset

2) Priors arranged by effective g-value (increasing top to bottom)

Table 9a
Posterior Inclusion Probabilities Across Parameter Priors
Simulated Data, Model1, k=15, n=50

	Priors Arranged By Effective g-Value (increasing left to right)										
Regressor	**11**	**9**	**6**	**1**	**12**	**3**	**4**	**2**	**8**	**5**	**7**
z_1	100	100	100	100	100	100	100	100	100	99.9	99.5
z_7	100	100	99.3	100	100	100	99.8	99.2	99.6	94.4	90.9
z_{11}	99.6	99.6	96.9	99.9	99.7	99.7	98.6	95.6	97.9	84.3	79
z_5	70	67	65.5	73.7	70.5	71.2	67.8	46.2	65.1	36.9	34.5
z_2	18.5	23.6	37.3	34.9	32.2	34.9	37	20.9	35.7	22.6	22.3
z_4	19.9	23.1	36.7	32.9	30.7	33.2	35.8	22.1	34.9	26	26.3
z_{14}	18.8	13.8	32.5	27.4	23.4	26.8	31.1	11.2	29.2	14.7	15.9
z_9	10.6	8.7	31.3	20	16.7	20.1	28.2	8.8	26.3	11.4	12.5
z_3	9	9.3	29.2	21.7	18.1	21.5	27.3	8.4	25.4	11.4	12.5
z_{13}	10.7	7.5	22.1	14.1	12.5	14.4	19.6	7.7	18.6	11	12.4
z_{12}	10.2	8.9	20.2	15	13.6	15.2	18.6	8.2	17.7	10.5	11.3
z_8	6.7	5.3	18.1	9.5	8.7	10.1	15.2	7.2	14.7	11.2	12.6
z_{15}	6.4	6.1	15.3	9.7	9.1	10.3	13.5	6.3	13.1	7.8	8.4
z_6	5.1	4.2	7.3	4.9	5.1	5.4	6.4	5.2	6.5	6.8	7.2
z_{10}	5.2	4.4	7.1	4.9	5.2	5.4	6.3	5.3	6.4	7.1	7.5
# effects	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**3**	**4**	**4**	**4**

Table 9b
Posterior Inclusion Probabilities Across Parameter Priors
Simulated Data, Model 1, k=15, n=100

	Priors Arranged By Effective g-Value (increasing left to right)										
Regressor	**11**	**9**	**1**	**12**	**6**	**3**	**2**	**4**	**8**	**5**	**7**
z_1	100	100	100	100	100	100	100	100	100	100	100
z_7	100	100	100	100	100	100	100	100	100	99.6	97.9
z_{11}	99.4	99.4	99.7	99.5	99.5	99.5	97.6	99.1	98.1	86.5	75.6
z_5	92.9	92.9	95.6	94.5	94.5	94.9	83.8	93.9	90.5	57.6	43.6
z_{15}	79.9	81.1	87.8	85	85.1	86.2	63.2	85.1	78.8	35.8	28.3
z_6	15.6	15.4	22.1	21.2	21.3	23.7	14.9	39	38	13	12.3
z_{12}	13.7	13.2	19.2	18.3	18.4	20.5	12.4	33.2	32.2	10.9	10.4
z_4	14.3	15.8	17.3	17.9	18	19.1	23	27.5	29.7	33.6	34.2
z_{13}	7.7	6.9	9.9	9.7	9.7	10.9	7.1	16.7	16.6	7.9	8.8
z_{10}	4.8	5.1	7.9	7.6	7.6	8.7	5.2	17.7	17.8	5.3	5.4
z_3	4	6.1	7.4	7.6	7.6	8.3	7.7	12.3	13.1	9.1	8.7
z_2	3.2	5	7	6.9	6.9	7.8	5.4	13.2	13.4	5.9	5.9
z_8	6	5.6	7	7	7.1	7.7	6.4	11	11.3	7.4	7.7
z_9	4.9	4.6	6.8	6.6	6.7	7.6	4.9	14.3	14.4	5.2	5.2
z_{14}	4.6	4.3	6	5.9	6	6.7	4.6	10.9	11.1	5	5.3
# effects	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**4**	**3**

Table 9c
Posterior Inclusion Probabilities Across Parameter Priors
Simulated Data, Model 2, k=40, n=100

	Priors Arranged By Effective g-Value (increasing left to right)										
Regressor	**11**	**9**	**1**	**12**	**6**	**3**	**4**	**8**	**2**	**5**	**7**
z_1	1.5	1.8	2.8	2.4	2	2.7	0.8	1.3	0.8	2.1	2
z_2	0.9	1.2	8.6	1.7	1.5	2	0.2	0.1	0	0	0
z_3	4.1	4.8	13.9	4.9	4.5	5.6	0.4	0.2	0	0.4	0.9
z_4	0.6	0.6	1.6	1.1	1.3	1.2	0.1	0	0	1	2.1
z_5	0.3	0.4	1.9	0.8	0.5	0.9	0.2	0.6	0.7	0.2	0.1
z_6	0.4	0.5	3.9	1	0.5	1.1	0.1	0	0	0	0
z_7	0.3	0.3	1.5	0.8	0.1	0.9	0.2	0.5	0.9	0.5	0.3
z_8	0.4	0.6	4.5	1	0.1	1.1	0.1	0.1	0.1	1	1.1
z_9	0.3	0.4	2.5	0.8	0.5	0.9	0.1	0	0	0	0
z_{10}	0.4	0.4	1.6	0.9	0.6	0.9	0.1	0	0.1	1.3	1.9
z_{11}	6.1	6.7	14.3	6.1	6.2	6.8	0.5	0.2	1.2	6.3	7
z_{12}	10.7	14.2	33.2	11.7	10.7	13.2	1.8	0.7	0	0	0
z_{13}	0.3	0.4	3	0.9	0.6	1	0.1	0	0	0	0.2
z_{14}	12.7	12.6	6.8	15.7	14.7	16	12	7.8	0.5	0.4	0.2
z_{15}	0.4	0.5	3.9	0.9	0.1	1.1	0.1	0	0	0	0.1
z_{16}	1.5	1.8	4.9	2.1	2.3	2.4	0.2	0.1	0	0.6	1.2
z_{17}	0.5	0.6	2.5	1	1	1.1	0.2	0.4	0.4	2.6	3.4
z_{18}	10.4	10.6	7.1	8.8	9.6	9.3	14.7	22.4	29.9	23.7	17.8
z_{19}	0.8	1	6.1	1.4	1.3	1.7	1.4	3.6	9.3	10.6	9
z_{20}	0.6	0.7	2.7	1.2	1.4	1.3	1.7	1.5	1.9	1.2	1.2
z_{21}	4.4	7	57.1	4.2	4	5.3	0.4	0.9	2.1	1	0.6
z_{30}	35.3	41.9	94	26.5	26.5	30	3.8	1.5	0	0.4	1.1
z_{38}	44.6	50.9	95.9	38.4	38.4	41.2	20.1	11.9	1.3	0.6	0.3
z_{33}	98.7	99	100	93.3	93.2	93.7	38.2	19.8	0.5	3.8	5.3
z_{22}	72.2	75.4	98.6	50.9	49.7	54.8	7.4	9.1	21.8	40.4	45.2
z_{25}	99.7	99.8	100	96.8	96.6	97.1	29.1	14.7	1.1	1.1	0.8
z_{27}	100	100	100	99.3	99.4	99.3	64.5	39.4	0.9	0.3	0.3
z_{32}	99	99.3	100	94.2	93.8	94.7	50.8	30.9	1.3	1	1.4
z_{35}	100	100	100	100	100	100	72.5	45.7	2.6	2.7	2.9
z_{23}	100	100	100	100	100	100	81.9	56.9	3.1	2	2.3
z_{37}	100	100	100	100	100	100	83.1	57.8	4.6	1.6	1.1
z_{39}	100	100	100	100	100	100	97.3	86.7	31	13.4	10.4
z_{31}	100	100	100	99.4	99.5	99.4	77.4	79	67.6	45.7	35.2
z_{29}	100	100	100	100	100	100	99.9	98.7	78.3	37.3	24.6
z_{24}	100	100	100	100	100	100	99.4	95	55.7	26.6	19.9
z_{36}	100	100	100	100	100	100	80.7	61.4	28.4	19.6	14
z_{28}	100	100	100	100	100	100	99.9	99.2	90.2	64.5	50.5
z_{26}	99	99.2	100	92.7	93.5	93.3	55.8	66.7	82.2	85.6	86.1
z_{40}	100	100	100	99.5	99.4	99.5	85.1	89.3	100	95.3	86.8
# effects	**16**	**17**	**19**	**16**	**15**	**16**	**13**	**10**	**7**	**4**	**3**

1) Light brown shaded variables should have an effect
2) Dark grey shaded cells indicated posterior inclusion probability over 50% (Jeffreys, 1961)
3) Priors 9 and 10 are identical in the simulated datasets
4) Uniform model priors through

Table 10a
Predictive Performance
***Relative* to Parameter Prior 1**
Model 1, k=15 n=50
(Uniform Model Prior)
Trials: 400

Prior	Median[a]	+ / 100[b]	Significance[c]
		MSE	
11	0.003	55	0.020
9	0.002	56	0.007
6	0.000	52	0.291
12	0.001	55	0.040
3	0.000	52	0.291
4	0.007	59	0.000
2	0.010	60	0.000
8	0.008	59	0.000
5	0.054	75	0.000
7	0.088	79	0.000
		CRPS	
11	0.007	72	0.000
9	0.004	69	0.000
12	0.000	47	0.876
6	-0.001	42	0.999
3	-0.001	42	0.999
4	-0.002	41	1.000
8	-0.001	42	0.999
2	0.001	54	0.060
5	0.001	54	0.073
7	0.002	57	0.003
		LPS	
9	0.053	55	0.040
8	0.817	71	0.000
11	0.057	56	0.016
6	0.049	56	0.007
3	0.049	56	0.007
2	0.849	70	0.000
4	0.768	71	0.000
7	2.962	87	0.000
5	2.330	84	0.000
6	1.542	76	0.000

Table 10b
Predictive Performance
***Relative* to Parameter Prior 1**
Model 1, k=15 n=100
(Uniform Model Prior)
Trials: 400

Median[a]	+ / 100[b]	Significance[c]
	MSE	
0.002	66	0.000
0.014	69	0.000
0.000	59	0.000
0.002	67	0.000
0.000	59	0.000
0.006	64	0.000
0.006	66	0.000
0.014	69	0.000
0.039	84	0.000
0.074	91	0.000
	CRPS	
0.007	319	0.000
0.008	82	0.000
0.000	46	0.939
-0.003	26	1.000
-0.003	26	1.000
-0.006	23	1.000
-0.006	25	1.000
-0.001	48	0.773
-0.004	33	1.000
-0.004	38	1.000
	LPS	
0.120	71	0.000
2.257	77	0.000
0.142	70	0.000
0.162	56	0.009
0.162	56	0.009
1.233	70	0.000
1.468	69	0.000
5.801	94	0.000
4.043	90	0.000
3.105	85	0.000

Table 10c
Predictive Performance
***Relative* to Parameter Prior 1**
Model 2, k=40 n=100
(Uniform Model Prior)
Trials: 400

Median[a]	+ / 100[b]	Significance[c]
	MSE	
0.007	90	0.000
0.006	90	0.000
0.001	64	0.000
0.003	79	0.000
0.001	64	0.000
0.007	62	0.000
0.022	84	0.000
0.017	71	0.000
0.059	88	0.000
0.097	93	0.000
	CRPS	
-0.002	47	0.927
0.000	50	0.520
0.001	54	0.073
0.002	57	0.005
0.002	57	0.005
0.012	71	0.000
0.013	73	0.000
0.004	58	0.001
0.013	71	0.000
0.016	74	0.000
	LPS	
0.331	79	0.000
1.872	73	0.000
0.463	81	0.000
-0.092	45	0.988
-0.092	44	0.988
1.475	76	0.000
0.954	67	0.000
6.124	92	0.000
4.274	87	0.000
2.734	77	0.000

[a] Median refers to the median inprovement in the score attained by the UIP compared to a given alternative prior

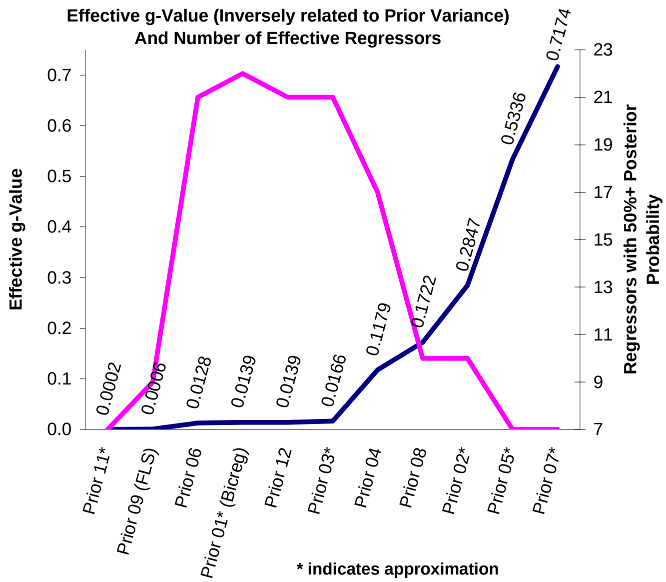
[b] Number of successes per 100 trials where "success" is a better predictive score by the UIP than by the alternative prior

[c] Significance refers to the binomial p values P(X > or = z) for the given number of trials and successes; where success is

1) Priors 9 and 10 are identical in the simulated dataset

2) Priors arranged by effective g-value (increasing top to bottom)

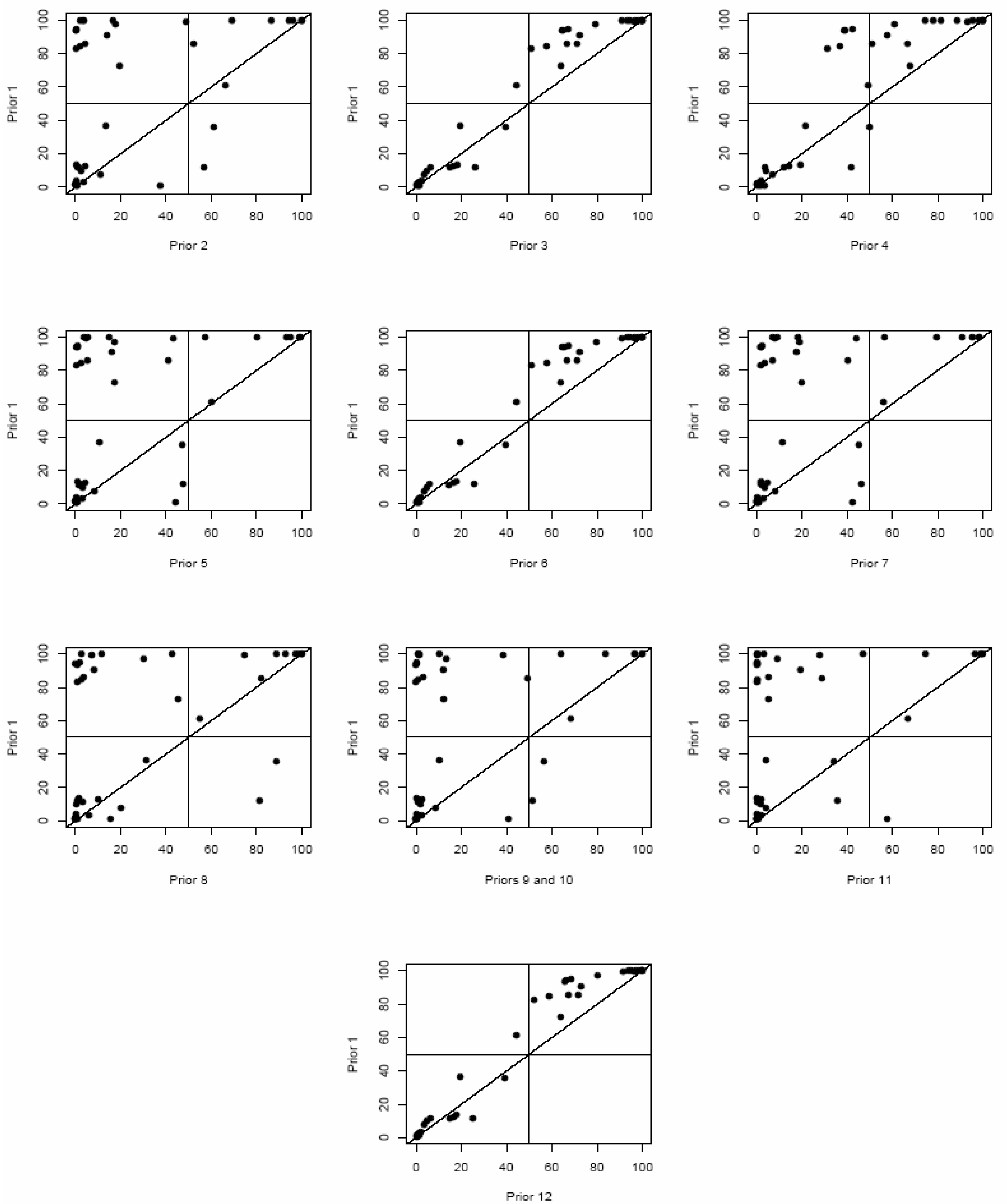
Figure 1



1) When priors depend on the exact model size, p_k, Figure 1 approximates the prior using the expected model size. Priors 11 and 1 are not exact g priors, so the g value is also an approximation

2) Priors 9 and 10 are identical in the growth context

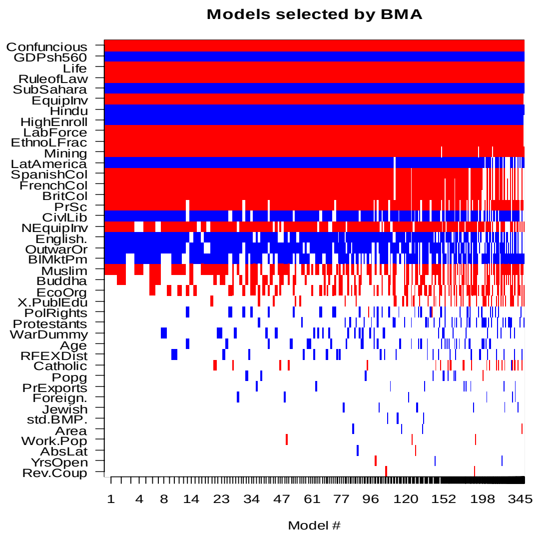
Figure 2
Correlation of Posterior Inclusion Probabilities Across Parameter Priors
(Growth Dataset)



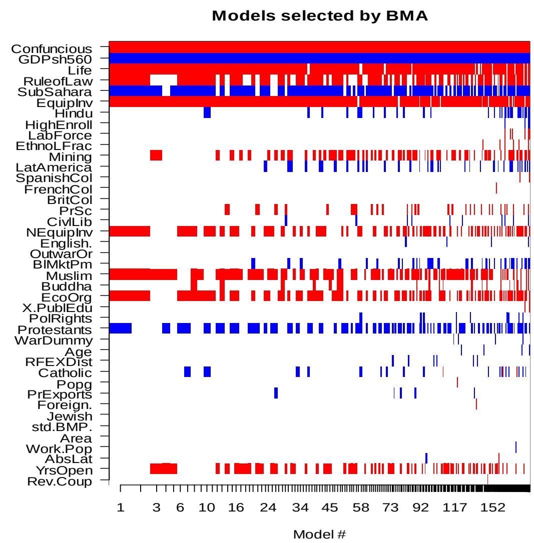
1) Priors 9 and 10 are identical in the growth context

Figure 3
Regressors Included in Best Models

a) Prior 1 (uniform model prior)



b) Prior 9 (uniform model prior)



Priors 9 and 10 are identical in the growth context.

c) Prior 1 (prior model size = 7)

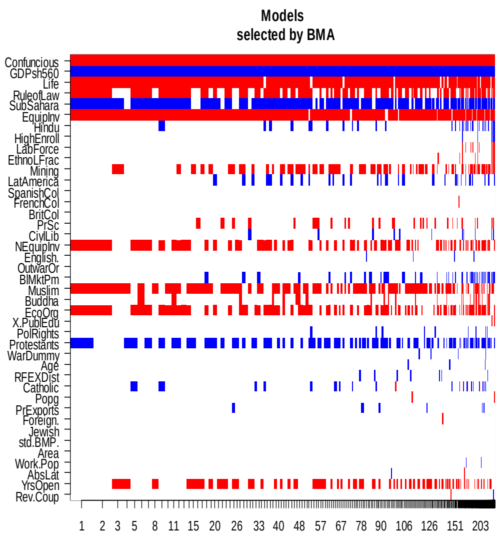


Figure 4
Posterior Means Across Parameter Priors

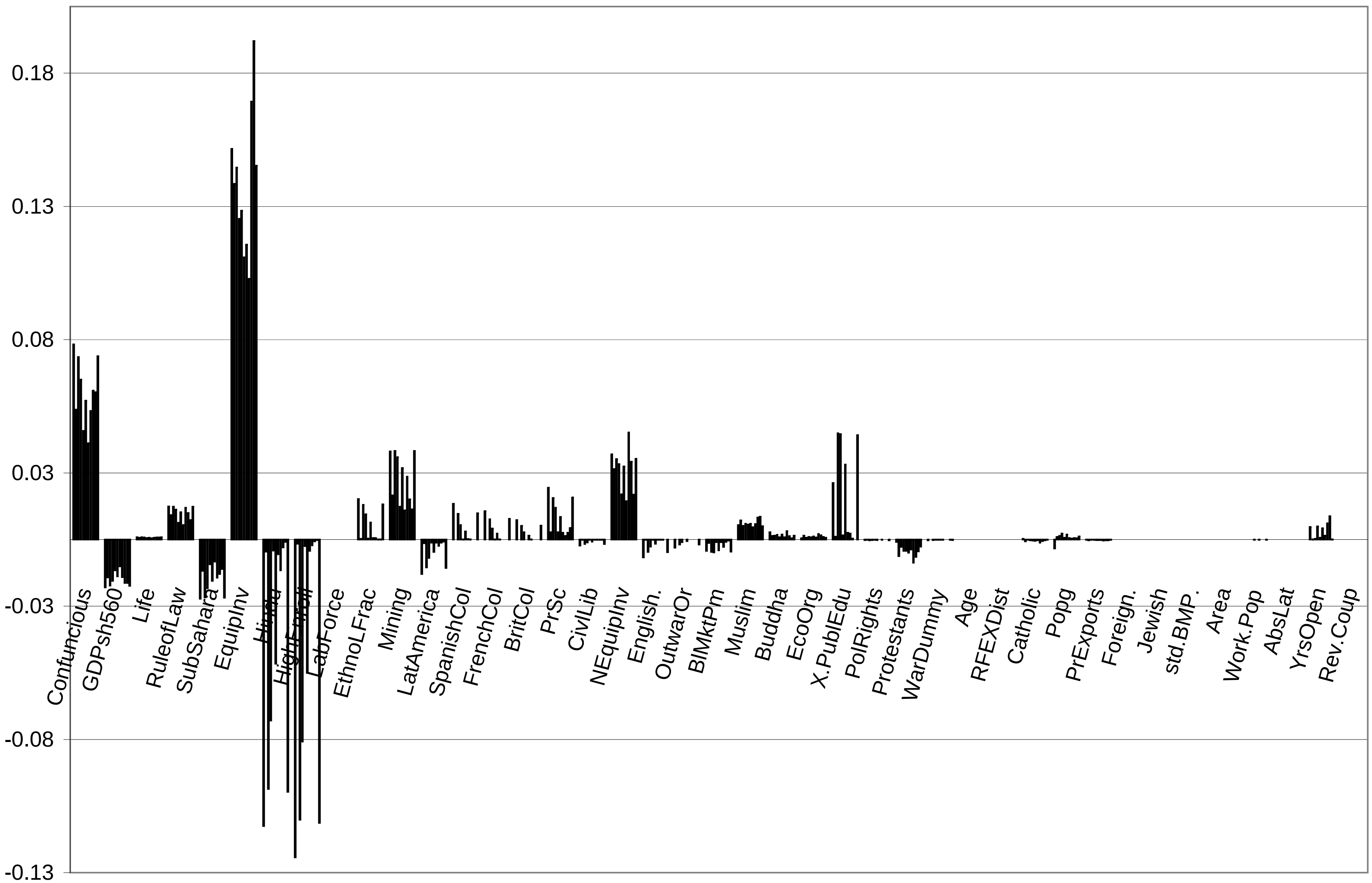


Figure 5a
Effective g-Value (Inversely Related to Prior Variance)
And Number of Effective Regressors (Posterior > 50%)
Simulated Data, Model, 1 k=15, n=50

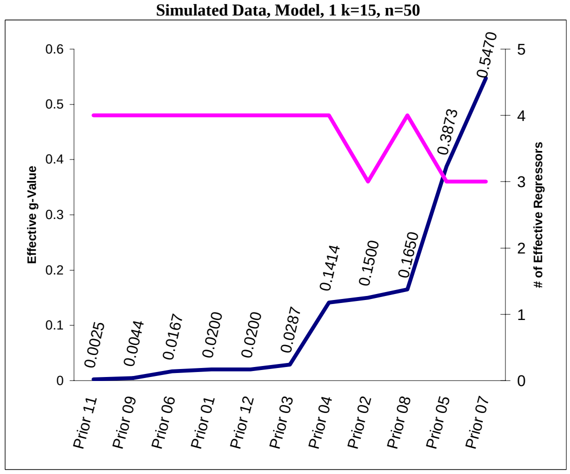


Figure 5b
Effective g-Value (Inversely Related to Prior Variance)
And Number of Effective Regressors (Posterior > 50%)
Simulated Data, Model 1, k=15, n=100

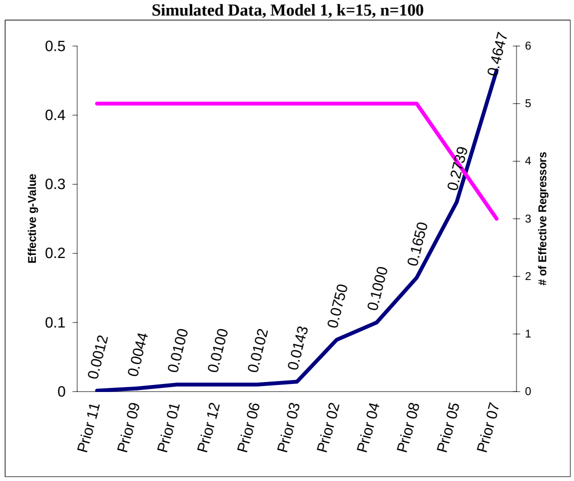
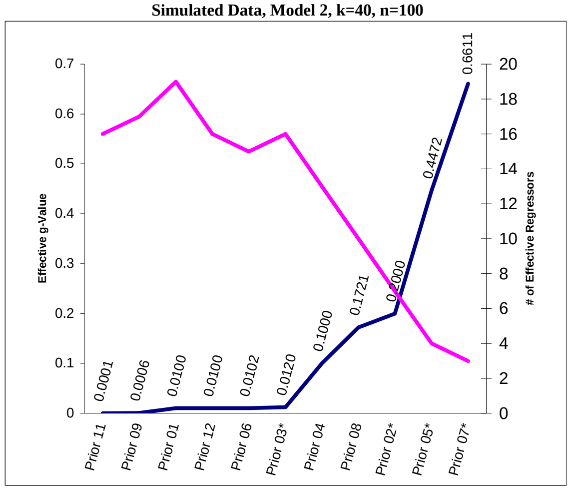


Figure 5c
Effective g-Value (Inversely Related to Prior Variance)
And Number of Effective Regressors (Posterior > 50%)
Simulated Data, Model 2, k=40, n=100



1) Priors 9 and 10 are identical in the simulated datasets
2) Priors 1 and 12 have the same g-value
3) Priors arranged by effective g-value (increasing left to right)