

Robust Growth Determinants

Gernot Doppelhofer and Melvyn Weeks

24 January 2011

CWPE 1117

Robust Growth Determinants*

Gernot Doppelhofer[†]
NHH and CESifo

Melvyn Weeks[‡]
University of Cambridge

January 24, 2011

Abstract

This paper investigates the robustness of determinants of economic growth in the presence of model uncertainty, parameter heterogeneity and outliers. The robust model averaging approach introduced in the paper uses a flexible and parsimonious mixture modeling that allows for fat-tailed errors compared to the normal benchmark case. Applying robust model averaging to growth determinants, the paper finds that eight of eighteen variables found to be significantly related to economic growth by Sala-i-Martin *et al.* (2004) are sensitive to deviations from benchmark model averaging. For example, the GDP shares of mining or government consumption, are no longer robust or economically significant once deviations from the normal benchmark assumptions are allowed. The paper identifies outlying observations – most notably Botswana – in explaining economic growth in a cross-section of countries.

Keywords: Determinants of Economic Growth, Robust Model Averaging, Heteroscedasticity, Outliers, Mixture models.

JEL Classifications: C11, C21, C52, O20, O47, O50

*We thank Jesus Crespo-Cuaresma, Nils Gottfries, Chris Papageorgiou, Hashem Pesaran, Xavier Sala-i-Martin, Jan Magnus, Pablo Selaya, Jan-Egbert Sturm, Andrew Warner, Arnold Zellner, seminar participants at Aberdeen, BI Oslo, Birkbeck, Cambridge, CESifo Munich, ETH Zurich, Innsbruck, NHH, Norges Bank, Rome and IHS Vienna and conference participants at the EEA 2010 meeting, the ES 2009 Winter meeting, and the Nordic Summer Symposium in Macroeconomics (NORMAC) 2010 for helpful comments. Doppelhofer thanks KOF at ETH Zurich and the economics department at UCL for their hospitality. All errors are our own.

[†]Department of Economics, Norwegian School of Economics and Business Administration (NHH), Helleveien 30, N-5045 Bergen. Email: Gernot.Doppelhofer@nhh.no, Tel: +47 55959360, Fax: +47 55959543.

[‡]Faculty of Economics and Clare College, University of Cambridge, Cambridge CB3 9DD, UK. Email: mw217@cam.ac.uk.

Contents

1	Outliers and Robustness	3
1.1	Alternative Robust Approaches	4
2	Benchmark Model Averaging	6
2.1	Model Space Prior	8
2.2	Estimation of Posterior Objects	9
3	Robust Model Averaging	11
3.1	Scale Mixtures of Normals	11
3.2	Estimation of Robust Posterior Objects	13
4	Robustness of Growth Determinants	15
4.1	Data and Benchmark Results	15
4.2	Main Robust Findings	16
4.3	Sensitivity Analysis	18
5	Conclusion	20
	References	21
A	Computational Appendix	25
A.1	Random and Stratified Sampling	25
A.2	Gibbs Sampling	25
A.3	Numerical Convergence Criteria	27
B	Data Appendix	28

[A]ny good approach to model uncertainty should ideally be robust to observations that are measured with error, or drawn from a different regime.

Temple (2000)

The empirical analysis of determinants of economic growth has generated a large literature among economists and applied researchers. This literature is plagued by a number of problems that can have important consequences for statistical inference and economic implications.¹ In particular, Temple (2000) highlights three key issues: (i) model uncertainty, (ii) parameter heterogeneity, and (iii) outliers. This paper investigates the robustness of inference about growth determinants to these specification problems.

Consider first model uncertainty about the process driving economic growth. A challenging problem is the large number of potential theories of economic growth combined with a limited number of observations. Brock and Durlauf (2001) refer to the “open-endedness” of economic growth theories, in the sense that the validity of one theoretical model does not contradict the validity of other models. A recent and quickly growing literature employs model averaging techniques² to address model uncertainty and the effect on inference and policy analysis on growth determinants. Early papers that address model uncertainty in growth regressions include Fernandez, Ley and Steel (2001a), and Sala-i-Martin, Doppelhofer and Miller (2004). Recently, model averaging has been applied in the context of growth empirics to investigate the sensitivity to prior information (Ley and Steel, 2009), and predictive performance (Eicher, Papageorgiou and Raftery, 2009).

A second issue is parameter heterogeneity, which is quite plausible when analyzing economic growth across countries. Studies which have considered uncertainty due to parameter heterogeneity in growth empirics include Durlauf and Johnson (1995), Brock and Durlauf (2001), an entire special issue of the *Journal of Macroeconomics* edited by Papageorgiou (2007), Masanjala and Papageorgiou (2008) and Tan (2009). Even though these papers emphasize various sources of heterogeneity in the growth process, they utilize prior knowledge and condition on a particular mechanism thought to generate heterogeneity in the distribution of coefficients of interest. Statistical inference and economic implications in these papers are therefore conditional on correctly specifying the source of parameter heterogeneity. Other misspecification issues are typically ignored, and the models are typically assuming independent normal sampling and homoscedastic errors.

A third important problem is the presence of outliers in the data describing economic growth across countries. Schultz (1999, p. 71) notes that “Macroeconomic studies of growth often seek to explain differences in economic growth rates across countries in terms of [several variables]. However, these estimates are plagued by measurement

¹Durlauf, Johnson and Temple (2008) give an excellent critical survey over the econometric methods and challenges in empirical research on economic growth.

²For recent surveys see Hoeting, Madigan, Raftery and Volinsky (1999) or Doppelhofer (2008).

error and specification problems.” Deaton (2010) critically discusses measures of income across countries and over time. In a recent paper, Ciccone and Jarocinski (2010) investigate the sensitivity of inference on growth determinants when using different vintages of the Penn World Tables. Despite these data and specification problems, Zaman, Rousseeuw and Orhan (2001) note that a remarkably low number of papers address issues of robustness in the economic literature.

This paper contributes to the existing literature by dealing with model uncertainty and allowing for heterogeneity of unknown form, generated either by outliers or neglected parameter heterogeneity. The robust model averaging approach introduced in this paper combines model averaging with a flexible and parsimonious mixture modeling that allows for fat-tailed errors compared to the normal benchmark case. In particular, we assume that the regression errors in each model are drawn from an independent mixture normal distribution, scaling the error variances for each observation (country). The robust model averaging approach helps to make statistical and economic inference about growth determinants robust to the specifications problems discussed above.

The paper investigates the robustness of growth determinants by applying robust model averaging to the Sala-i-Martin *et al.* (2004) data with 67 explanatory variables for the average growth rate of GDP per capita among 88 countries.³ Compared with the 18 variables that were found to be significantly related to economic growth using benchmark model averaging, we find that eight are no longer robust or economically significant in explaining economic growth once we allow for deviations from the normal benchmark assumptions. Examples of such non-robust growth determinants include the GDP shares of mining and government consumption and several regional variables. Our robust model averaging approach identifies several outliers in models explaining economic growth in a cross-section of countries. Several countries in Sub-Saharan Africa – Botswana, Central African Republic, Gabon, Zaire and Zambia – and the Philippines have variances between two to almost four times as large as under normal benchmark assumptions. This paper finds that the results in existing studies of economic growth determinants are thus not robust to allowing combinations of uncertainty about economic models, heterogenous parameters and outliers in the data.

The remainder of the paper is organized as follows: Section 1 discusses outliers and robustness of statistical inference and economic analysis, first conditional on a particular model, and then combined with model uncertainty. Section 2 reviews what we call benchmark model averaging, where model uncertainty is equivalent to the problem of variable selection. The robust model averaging approach introduced in section 3 combines both model uncertainty and a flexible mixture modeling of heteroscedastic errors. Section 4 applies robust model averaging to determinants of economic growth, and section 5 concludes.

³See the Data Appendix B for a list of variables and data summary.

1 Outliers and Robustness

Outliers due to measurement errors in the data or specification problems is a major problem encountered in the empirical growth literature. In a critical survey, Brock and Durlauf (2001) argue that much empirical work in this area suffers from “incredible” assumptions that are difficult to defend. More specifically, Brock and Durlauf warn that models of economic growth are unlikely to satisfy the basic assumption of exchangeability.⁴ Intuitively, the exchangeability assumption implies that the model explaining economic growth is the same across different countries and also selected subgroups of countries. For example, one specific requirement for exchangeability is that the model errors have the same variance across all observation points. When this is not the case, the errors are said to be heteroscedastic.

If specification problems are known *a priori*, it would be straightforward to adapt the growth model accordingly. In the more realistic case when the precise form of model misspecification are not known beforehand, inference should be conducted in a robust manner. One possible approach to modeling outlying observations is the introduction of *mean-shifts*. For example, Hendry and Santos (2005) propose to saturate the regression model by introducing a large number of dummy variables. Treating each outlying observation differently is problematic in the context of data limitations and model uncertainty in the empirical growth literature. Furthermore, economic theory offers little guidance about the appropriate form of parameter heterogeneity.

This paper uses instead the *variance-inflation* approach to accommodate outliers and robustify inference against *unknown* aberrant observations. The variance-inflation model has the advantage of being parsimonious and flexible, which makes it attractive given the numerous specification and data problems that plague the empirical growth literature. The variance-inflation approach uses a parsimonious mixture distribution that requires only one or two additional parameters. As an example, consider a combination of two distributions, with low and high variance, and within these distributions observations are identically and independently distributed. Combining these two gives a mixture distribution with different variances. Heterogeneous parameters can be handled through mixture distributions over one or more parameters (random coefficients model), but this would quickly get cumbersome with a large number of parameters. Fernandez and Steel (2000) examine Bayesian inference within the confines of the linear regression model, focussing on the theoretical basis of independent sampling from a scale mixture of normal distributions of the regression errors. A flexible approach to robust estimation becomes all the more important in the presence of model uncertainty, when theory gives us little guidance about the correct model.

The *robust model averaging* approach proposed in this paper simultaneously deals with model uncertainty and heterogeneity of unknown form. Inference and economic

⁴Brock and Durlauf (2001) refer to the *de Finetti Representation Theorem* that states that a sequence of exchangeable random variables (the model errors in the context of growth regressions) can be written as if generated by a mixture of identically and independently distributed (*i.i.d.*) random variable.

analysis are made robust with respect to outliers and unequal variances by allowing *a priori* for thicker tails of the distribution of regression errors compared to normal benchmark model averaging. Suppose that the regressions errors are independently normally distributed:

$$\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{\Omega})$$

with diagonal covariance matrix $\mathbf{\Omega} \equiv \text{diag}(\omega_1, \dots, \omega_N)$. The errors for each observation (or country $i = 1, \dots, N$) are scaled by a variance inflation term ω_i with an independent mixing distribution. Geweke (1993) demonstrates the equivalence of such a normal mixture model with an independent Chi-square prior to a model where errors are drawn from an independent Student- t distribution, where the degrees of freedom determine the fatness of the tails and the prior weight on outliers. In the context of model uncertainty about potential growth determinants, it is our goal to accommodate observations with different degrees of reliance within a framework that accommodates model uncertainty. Conditional on having estimated the covariance matrix $\mathbf{\Omega}$, all other quantities of interest for a given model are estimated by using Generalized Least Squares (GLS) instead of Ordinary Least Squares (OLS). The parameters of the model are drawn from their respective conditional distributions using the Gibbs sampler (see the Computational Appendix A for details).

In the following, we contrast the robust model averaging approach adopted in this paper with alternative ways to model outliers and unequal error variances.

1.1 Alternative Robust Approaches

The simplest method to incorporate a measure of robustness case is the application of a known (monotonic) transformation of the data. For example, a simple logarithmic transform may be applied if the problem is scale related, accounting for the observation that series that are growing exponentially often appear to have increasing variability. In the empirical investigation of growth determinants in section 4, we consider explanatory variables that have been suggested by the literature. For example, *Initial Income* is measured in natural logarithms, and many other variables are usually defined as ratios to GDP (see the Data Appendix B for a list of variables).

When dealing with outliers and misspecification, empirical researchers were initially restricted by limitations in statistical techniques and computing resources. Leamer (1982) and Leamer and Leonard (1983) propose so-called extreme bounds analysis (EBA) to test the sensitivity of parameters of interest to changes in the set of alternative models, represented by different combinations of additional control variables. Sturm and de Haan (2005) apply a version of EBA that uses re-weighted least squares first developed by Rousseeuw (1984). Zaman et al. (2001) highlight the importance of robust inference when applying least trimmed squares to a simple growth regression. Interestingly, even though our paper uses a larger sample and many more explanatory variables, we agree with Zaman *et al.* that Zambia is an important outliers when analyzing economic growth. Trimming of extreme observations can be interpreted as a

special case of the approach used in this paper, where observations identified as outliers receive zero weight.

Growth processes may also exhibit heterogeneity across a spatial dimension. LeSage and Parent (2007) allow for model uncertainty in spatial econometric models with an explicit treatment of spatial dependence of errors. Canova (2004) tests for the presence of convergence clubs among European regions, allowing for heterogeneous parameters across particular subgroups. Corrado, Martin and Weeks (2005) test for regional convergence clusters across Europe. Crespo-Cuaresma, Doppelhofer and Feldkircher (2009) investigate the determinants of economic growth in European regions, allowing for spatial spillovers across regions. The robust model averaging approach can also account for spatial clustering of errors by accommodating outliers and heteroscedastic errors.

In both Classical and Bayesian settings estimates will be sensitive to the particular set of assumptions which underlie the approach. For example, our approach to robustness, and that of alternative approaches based on mixture modeling, depends on parametric (prior) assumptions about the distribution of errors. An alternative approach to robust inference is to use the Bayesian bootstrap developed by Rubin (1981).⁵ The logic of the Bayesian bootstrap is to consider parameters as functionals of the data (i.e. moment condition) and to sample directly from the posterior distribution of the data. Poirier (2008) shows that posterior weights can then be used to weight individual observations such that the resulting parameter estimates have a weighted least squares representation. Another flexible approach is the use of heteroscedasticity-consistent standard errors (HCSE). Even though they are biased in finite samples, HCSE represent an improvement upon OLS estimates using a minimal set of assumptions (see White, 1980; MacKinnon and White, 1985). Lancaster (2009) demonstrates that HCSE are reasonable approximations to the posterior standard deviations around the OLS estimator (see also Poirier, 2008). Section 4.3 investigates the sensitivity of inference to using HCSE as alternative to our approach.

Despite the emergence of this literature it appears that much of the advances in robust methods in Bayesian inference have been confined to single models. A notable exception is the work of Hoeting, Raftery, and Madigan (1996), who develop an approach that simultaneously accounts for model uncertainty and outlier identification by introducing a prior for the proportion of outlying observations. Recently, Gottardo and Raftery (2007) adopt a unifying approach to Bayesian robust variable and transformation selection. Magnus, Wan and Zhang (2010) use a version of the weighted average least squares (or WALS) estimator with nonspherical disturbances in an analysis of the Hong Kong housing market. Our paper also addresses the problem of model uncertainty and outlier detection within a unifying framework. Following Geweke (1993), we introduce a parsimonious mixing distribution for the regression errors using both fixed and variable hyperparameters. The paper applies robust model averaging to determinants of economic growth and finds that the results are sensitive to outliers and neglected

⁵One variant of this approach, so-called *Bagging* (short for Bootstrap Aggregating) is used to generate robust predictions accounting for data problems (see Breiman, 1996; and Clyde and Lee, 2001).

heterogeneity.

2 Benchmark Model Averaging

Consider the typical cross-country growth regression of the form:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \quad (1)$$

where the $N \times 1$ vector \mathbf{y} contains observations on economic growth for N countries, the explanatory variables or regressors are written compactly as $N \times k$ matrix $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_k)$, and associated conformable vector of unknown slope parameters β . Note that an intercept is always included in the growth regression. The benchmark case usually assumes that the $N \times 1$ regression errors ε are (1) normally distributed, and (2) *conditionally* homoscedastic. The benchmark model averaging case discussed in this section takes these assumptions as given, but section 3 will return to this issue in more detail.

Suppose that a researcher wishes to assess the effect associated with a particular growth determinant, captured by its slope coefficient. A difficulty arises from the large number of potential explanatory variables, compared to the limited number of available observations N across countries.⁶ With K potential regressors, the model space \mathcal{M} is the set of all 2^K combinations of linear models. Each model M_j is described by a $k \times 1$ binary vector $\gamma = (\gamma_1, \dots, \gamma_K)'$, where a one (zero) indicates the inclusion (exclusion) of a variable \mathbf{x}_k in regression (1). Let \mathbf{X}_j be the set of regressors included in model M_j with associated slope coefficient β . The researcher can estimate its distribution $p(\beta|M_j)$ conditional on model M_j . The *unconditional* distribution of coefficients can be derived by integrating out all aspects of model uncertainty, including the space of models \mathcal{M} . A maintained assumption throughout is that the explanatory variables \mathbf{X} are predetermined (weakly exogenous) and independent of parameters β and σ . The posterior density $p(\beta|\mathbf{y})$ can then be expressed as function of sample observations of the dependent variable \mathbf{y} .

The unconditional posterior distribution of the slope coefficient β is given by

$$p(\beta|\mathbf{y}) = \sum_{j=1}^{2^K} p(\beta|M_j, \mathbf{y}) \cdot p(M_j|\mathbf{y}) \quad (2)$$

where $p(\beta|M_j, \mathbf{y})$ is the conditional distribution of β given model M_j . The posterior model probability $p(M_j|\mathbf{y})$ propagates model uncertainty into the posterior distribution of model parameters. By Bayes' rule, the posterior model probability can be written as

$$\begin{aligned} p(M_j|\mathbf{y}) &= \frac{l(\mathbf{y}|M_j) \cdot p(M_j)}{p(\mathbf{y})} \\ &\propto l(\mathbf{y}|M_j) \cdot p(M_j) \end{aligned} \quad (3)$$

⁶For example, Durlauf *et al.* (2008) list in the Appendix of their survey 145 different growth determinants that have been suggested in the literature. However, similar small sample problems can also be encountered in many other areas using non-experimental data.

such that the posterior model probability (weight) of model M_j is proportional to the product of the model-specific marginal likelihood $l(\mathbf{y}|M_j)$ and the prior model probability $p(M_j)$. The model weights are converted into probabilities by normalizing relative to the set of all 2^K models:

$$p(M_j|\mathbf{y}) = \frac{l(\mathbf{y}|M_j) \cdot p(M_j)}{\sum_{l=1}^{2^K} l(\mathbf{y}|M_l) \cdot p(M_l)} \quad (4)$$

We follow the (Bayesian) model averaging literature by assuming the following prior structure for parameters in each model. The prior slope coefficients β are normally distributed with mean zero and variance $\sigma^2 \mathbf{V}_{0j}$:

$$p(\beta|\sigma^2, M_j) \sim N(\mathbf{0}, \sigma^2 \mathbf{V}_{0j}) \quad (5)$$

For the prior error variance term σ^2 that is common to all models, we assume a non-informative (diffuse) prior that imposes a minimum of prior information. Alternatively, one could assume a proper, inverse-Gamma prior distribution for the error variance σ^2 , which is the natural conjugate prior for the normal regression model. We check for sensitivity of results using a proper Bayesian prior in section 4.3.

The prior variance matrix is assumed to be proportional to the sample covariance

$$\mathbf{V}_{0j} = (g_0 \mathbf{X}'_j \mathbf{X}_j)^{-1} \quad (6)$$

with factor of proportionality g_0 . This *g-prior* was first suggested by Zellner (1986), and is a convenient way to specify the prior variance matrix, in particular in the presence of considerable model uncertainty. Different values of the *g*-prior parameter g_0 have been proposed in the literature (see Fernandez, Ley and Steel, 2001b).⁷ We follow Sala-i-Martin *et al.* (2004) and assume that the prior distribution is dominated by the sample information, implying a diffuse prior variance.

The assumed prior structure introduces a minimum of prior information into the estimation. In the limit, when the sample information dominates the prior information, Leamer (1978) shows that the marginal likelihood of model M_j may be written as

$$l(\mathbf{y}|M_j) \propto N^{-k_j/2} \cdot SSE_j^{-N/2} \quad (7)$$

where k_j is the number of regressors and $SSE_j = (\mathbf{y} - \mathbf{X}_j \beta_j)'(\mathbf{y} - \mathbf{X}_j \beta_j)$ is the sum of squared errors in model M_j . The posterior model probability of model M_j is obtained by pre-multiplying (7) by the prior model probability $p(M_j)$ and dividing by the sum over all 2^K possible models:

$$p(M_j|\mathbf{y}) = \frac{p(M_j) \cdot N^{-k_j/2} \cdot SSE_j^{-N/2}}{\sum_{l=1}^{2^K} p(M_l) \cdot N^{-k_l/2} \cdot SSE_l^{-N/2}} \quad (8)$$

⁷Feldkircher and Zeugner (2009) warn that an overly diffuse prior concentrates estimation on a few models, what they call the "supermodel effect". This effect is contributing to the sensitivity of estimates across different samples of the Penn World Tables found by Ciccone and Jarocinski (2010).

The posterior model weights (8) equal the prior model weights times the (exponentiated) Bayesian Information Criterion (BIC), developed by Schwarz (1978). The BIC weights depend on the likelihood, through the term $SSE_j^{-N/2}$, but penalizes relatively large models through the penalty term $N^{-k_j/2}$. The implied preference for smaller models addresses to a certain extent collinearity among regressors. Explanatory variables that are very similar explain relatively less of the variation of the dependent variable which implies less weight on such models.

BIC model weights (8) have been extensively discussed in the literature. Alternative derivations include the so-called “unit information prior” discussed in Kass and Wassermann (1995), approximation to Bayes Factors by Kass and Raftery (1995) and Raftery (1995), benchmark priors by Fernandez, Ley and Steel (2001b), or the limiting case of a non-informative Jeffreys prior for the error variance with a particular choice of normalizing constant (Wasserman, 2000). Klein and Brown (1984) show that by minimizing the so-called Shannon information in the prior distribution, the BIC model weights (8) can be used in small samples. We adopt the BIC posterior model weights since they provide a reasonable approximation to proper Bayesian model weights and are consistent in large samples.

2.1 Model Space Prior

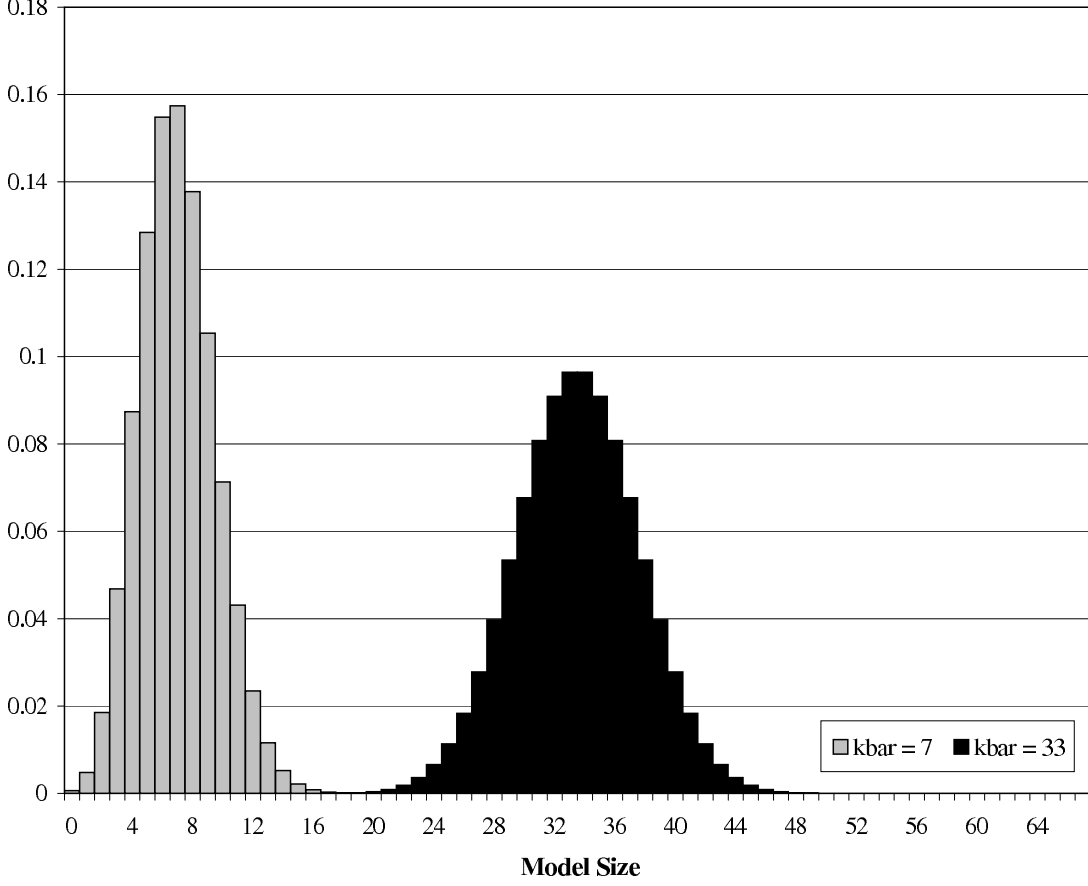
The last ingredient to make model averaging operative is the specification of a prior over the model space. Letting π_k be the independent prior inclusion probability of variable \mathbf{x}_k in model M_j , the probability for model M_j is given by the binomial distribution:

$$p(M_j) = \prod_{k=1}^K \pi_k^{\gamma_k} (1 - \pi_k)^{1-\gamma_k} \quad (9)$$

where the binary indicator variable γ_k measures inclusion (exclusion) of variable \mathbf{x}_k .⁸ One approach is to assume a completely diffuse or uniform prior across all models, which corresponds to a prior inclusion probability equal to $\pi_k = 1/2$ for all variables. However, with a relatively large number of regressors, a uniform prior implies that the great majority of prior probability is allocated to models with a large number of variables. As an alternative, Sala-i-Martin *et al.* (2004) advocate in their BACE approach a preference for more parsimonious models with a smaller prior expected model size $\bar{k} = 7$, which seems reasonable given the relatively large number of growth determinants ($K = 67$). Figure 1 contrasts the prior distribution over model of different sizes k for the BACE benchmark case with prior inclusion probability $\pi_k^{BACE} = \bar{k}/K = 0.104$ with the case of uniform model priors with $\pi_k^U = 0.5$. The two distributions clearly differ with regards to the assumed average size of the models ($\bar{k} = 7$ in the BACE case, compared to 33.5 under the uniform model prior).

⁸Mitchell and Beauchamp (1988) first suggested this prior with discrete probability mass or “spike” at zero, representing the prior uncertainty that a regressor should be included. George and McCulloch (1993) propose a Bayesian alternative of using a proper prior distributions with large variance.

Figure 1: Prior Probabilities by Model Size: Benchmark Case with Prior Model Size $\bar{k} = 7$ and Uniform Prior with $\bar{k} = 33$.



A disadvantage of assuming a fixed prior mean model size parameter is that researchers may differ in their opinions of what constitutes a reasonable prior model size. The Bayesian solution is to treat the expected mean model size as random variable. Following Brown, Vanucci and Fearn (1998), we introduce another layer of prior information that combines the independent Bernoulli sampling for each variable with a conjugate Beta hyper-prior for the binomial proportion parameter π_k (see also Ley and Steel, 2009). Section 4.3 investigates the sensitivity of our results to allowing for random model size.

2.2 Estimation of Posterior Objects

With all ingredients to calculate posterior model probabilities (3) in place, we can proceed to estimate the unconditional distribution quantities of interest that incorporates model uncertainty. For example, the unconditional distribution of the slope coefficient $p(\beta|\mathbf{y})$ is obtained by integrating over the space of models \mathcal{M} . A summary measure of

importance of an explanatory variable \mathbf{x}_k is its *posterior inclusion probability*

$$p(k|\mathbf{y}) = \sum_{j=1}^{2^K} \mathbf{1}(\gamma_k = 1|\mathbf{y}, M_j) \cdot p(M_j|\mathbf{y}) \quad (10)$$

which represent the unconditional probability that variable \mathbf{x}_k enters the regression model and is relevant in explaining the dependent variable (see Leamer, 1978; and Mitchell and Beauchamp, 1988). The posterior inclusion probability can be contrasted with the prior inclusion probability to see if after visiting the various models, the data increase or decrease out confidence that variable \mathbf{x}_k is important in explaining the dependent variable.

A researcher might be also interested in drawing inference about the economic importance of variables. The unconditional mean and variance can be calculated in a straightforward manner from their conditional (model specific) counterparts and posterior model probabilities (see Leamer, 1978). The *posterior mean* of the slope parameter β_k associated with variable \mathbf{x}_k is given by

$$E(\beta_k|\mathbf{y}) = \sum_{j=1}^{2^K} p(M_j|\mathbf{y}) \cdot \hat{\beta}_{kj} \quad (11)$$

where $\hat{\beta}_{kj} = E(\beta_k|\mathbf{y}, M_j)$ is the OLS estimate of the slope parameter β_k given model M_j . The *posterior variance* of slope β_k is given by

$$V(\beta_k|\mathbf{y}) = \sum_{j=1}^{2^K} p(M_j|\mathbf{y}) \cdot V(\beta_k|\mathbf{y}, M_j) + \sum_{j=1}^{2^K} p(M_j|\mathbf{y}) \cdot [\hat{\beta}_{kj} - E(\beta_k|\mathbf{y})]^2 \quad (12)$$

consists of two terms: the weighted sum of conditional (model-specific) variances and an additional term taking into account the difference between conditional and posterior estimates of mean coefficients. Given that we assume a prior structure that is dominated by sample information, conditional (model-specific) variances are estimated by the maximum likelihood estimator $V(\beta_k|\mathbf{y}, M_j) = \hat{\sigma}_j^2(\mathbf{X}_j'\mathbf{X}_j)^{-1}_{kk}$, with error variance estimate $\hat{\sigma}_j^2 \equiv SSE_j/(N - k_j)$.

To evaluate the economic effect of a variable, a researcher might be interested in looking at estimates of the mean and variance of the slope coefficients *conditional* on including a particular variable in the regression, but unconditional with respect to the model space \mathcal{M} and the inclusion of other explanatory variables. The *conditional posterior mean* for the slope coefficient β_k is obtained by dividing the unconditional posterior mean (11) by the posterior inclusion probability (10):

$$E(\beta_k|\gamma_k = 1, \mathbf{y}) = \frac{E(\beta_k|\mathbf{y})}{p(k|\mathbf{y})} \quad (13)$$

Similarly, the variance conditional on including variable \mathbf{x}_k is calculated from the unconditional posterior estimates of moments (11), (12), and the posterior inclusion probability (10):

$$V(\beta_k|\gamma_k = 1, \mathbf{y}) = \frac{V(\beta_k|\mathbf{y}) + [E(\beta_k|\mathbf{y})]^2}{p(k|\mathbf{y})} - [E(\beta_k|\gamma_k = 1, \mathbf{y})]^2 \quad (14)$$

The *conditional posterior standard deviation* is simply the square root of the conditional variance (14). To measure of economic significance of an explanatory variable conditional on its inclusion in the model, one can divide the conditional posterior mean by the conditional posterior standard deviation. Brock and Durlauf (2001) provide a decision-theoretic foundation for the use of such standardized coefficients when undertaking economic inference.

Note that all the posterior statistics presented in this section – inclusion probability and moments, such as mean and variances – are estimated unconditionally with respect to the model space and are thereby taking model uncertainty into account. However, the benchmark model averaging framework does not take into account other forms of specification uncertainty, such as the sensitivity of model weights and inference to outliers and the assumed prior distribution. This will be addressed in the following section.

3 Robust Model Averaging

This section introduces our robust model averaging framework that addresses *both* model uncertainty and heteroscedastic errors due to parameter heterogeneity and outliers. Robust model averaging uses a flexible mixture of distributions for the error terms to make inference and economic analysis robust to parameter heterogeneity and outliers. The approach is parsimonious, because mixture distributions are a straightforward extension of the benchmark case of homoscedastic, normal errors requiring an additional parameter determining the fatness of the tails of the error distribution and the relative weight on outliers. We first introduce scale mixtures of normal distributions⁹ conditional on a particular model, and then show how this robust approach extends to averaging across different models.

3.1 Scale Mixtures of Normals

Recall the typical linear cross-country regression (1) from section 2. A model M_j is characterized by a set of regressors \mathbf{X}_j and associated slope coefficients β_j . We start by examining robustness to outliers and heterogeneity conditional on a given model M_j ¹⁰.

A maintained assumption in the normal benchmark case is that regression errors are normally distributed with homoscedastic errors. A useful point of departure is to suppose that the observations y_i are described by a combination of two normal distributions

$$p(y_i|\beta, \sigma^2, \pi, \rho) = (1 - \pi) \cdot N(y_i|\beta, \sigma^2) + \pi \cdot N(y_i|\beta, \rho\sigma^2) \quad (15)$$

⁹For the growth application considered in this paper, the mixture-normal approach captures the heterogeneity in the data well. McLachlan and Peel (2000) discuss other mixture distributions.

¹⁰This can be contrasted with Lange, Little and Taylor (1989) when estimating a model by maximum likelihood.

where the mixture is governed by two parameters. The parameter π is used to identify a subset of observations as potential outliers, and conditional on this subset, the parameter ρ controls the degree of variance-inflation for the outlying observations. Hoeting *et al.* (1996) adopt this approach in a study which simultaneously selects regressors and identifies outliers. In the particular application of their paper, the parameters π and ρ are treated as fixed, with the proportion of outliers π chosen based upon the size of the dataset.

More generally, we consider mixture distributions of the form

$$p(y_i|\beta, \sigma^2, M_j) = \int_0^\infty \frac{\omega_i^{-1/2}}{(2\pi)^{1/2}\sigma} \cdot \exp\left[-\frac{(y_i - \mathbf{X}_{ij}'\beta_j)^2}{2\sigma^2\omega_i}\right] \cdot p(\omega_i|\tau)d\tau, \quad (16)$$

which consists of two parts. The first term is a kernel of a normal distribution for the errors of a linear regression model for model M_j , *conditional* on the variance-inflation term ω_i . The second term adds a mixing distribution $p(\omega_i|\tau)$, described by an unknown parameter vector τ . The specification of the mixing distribution is equivalent to choosing a prior specification on the error variances ω_i specific to each observation, $i = 1, \dots, N$.

The mixing distribution $p(\omega_i|\tau)$ may be chosen on a number of grounds, including an analytically convenient form (conjugacy). A particularly convenient form of mixing distribution $p(\omega_i|\tau)$ is the Gamma-family, with two parameters, say a and b , which determine the shape and scale of the distribution. The conjugacy of the Normal-Gamma case is such that the conditional distribution of each of the variance terms ω_i is also Gamma, which is convenient for computation. Geweke (1993) shows that the normal mixture model with an independence prior for error variances across observations i

$$v/\omega_i \sim \chi^2(v), \quad i = 1, \dots, N \quad (17)$$

is equivalent to a model with independent Student- t errors with v degrees of freedom. We note that this distribution is also part of the Gamma-family since $\Gamma(a = v/2, b = 2)$ is equivalent to a Chi-squared distribution, $\chi^2(v)$. In this case the mixing distribution is controlled by a single parameter v controlling the degrees of freedom. Lower values of v imply a more skewed distribution with a higher probability of outliers and relatively larger variances. High values of v on the other hand imply errors drawn from a distribution close to the homoscedastic normal benchmark case described in section 2. Intuitively, the degrees of freedom v determine the fatness of the tails of the Student- t distribution and the prior weight on outliers. Lange *et al.* (1989) justify the use of student- t distributions to robustify statistical inference for a number of applications.

Outliers, neglected heterogeneity and other specification problems can result in heteroscedastic errors, and, as argued in section 1, these issues are particularly important in the empirical growth literature. To deal with these problems, the robust model averaging approach proposed in this paper assumes independently distributed, but heteroscedastic errors, $\varepsilon \sim N(\mathbf{0}, \sigma^2\Omega)$. The covariance matrix is assumed diagonal,

$\Omega \equiv \text{diag}(\omega_1, \dots, \omega_N)$, with independent variances ω_i for observations $i = 1, \dots, N$.¹¹ Following Geweke (1993), we assume an independent prior distribution (17) for the error variances, where the prior parameter v determines the fatness of the distribution tails or equivalently the prior weight on outliers. As the degrees of freedom become large ($v \rightarrow \infty$), the robust approach behaves like the limiting benchmark normal case, where regression errors become conditionally homoscedastic with covariance matrix equal the identity matrix.¹²

For the degrees of freedom parameter v we consider two alternative priors:

Case (a): Fixed degrees of freedom – We set the degrees of freedom parameter to fixed value $v = \bar{v}$. In the application to growth determinants in section 4, we consider the following values of $\bar{v} = 5, 25, 100$, where lower (higher) values correspond to fatter (thinner) tails of the error distribution, or a researcher’s prior for more (less) outliers.

Case (b): Random degrees of freedom – Alternatively, the degrees of freedom parameter are treated as random, and we assume an analytically convenient exponential distribution with prior mean v_0 :

$$p(v) \sim \exp(v|v_0) \quad (18)$$

The posterior distribution of the degrees of freedom v is given by $p(v|\omega_i, M_j) \sim p(\omega_i|v, M_j) \cdot p(v)$.

The next subsection shows how we can estimate all other objects of interest – slope coefficients and error variances – by drawing from their respective conditional distributions for each model M_j using the Gibbs sampler. The unconditional estimates are obtained by averaging across models in the same manner as described in section 2.

3.2 Estimation of Robust Posterior Objects

The prior assumptions over model parameters and the model space imply a convenient hierarchical structure, which leads naturally to estimation of robust posterior distributions for model parameters using the Gibbs sampler. Conditional on each model M_j , we draw quantities of interest from their conditional posterior distributions (for details, see the Computational Appendix A). Under mild conditions the Markov chain generated by the iterations of the Gibbs sampler converges to the full posterior distribution (see Chib, 2001).

¹¹Particular clustering of variances could be introduced *a priori*, but we focus on variance heterogeneity originating in the sample.

¹²When we allow for outliers and fat-tailed distribution of errors, posterior variances are only well defined for degrees of freedom exceeding a lower bound $v > 4$ (see Geweke, 1993). When applying robust model averaging to growth determinants in section 4, posterior values of v are around 20, pointing to sizeable heteroscedasticity, but clearly exceeding the lower bound for variances to exist.

The posterior distribution of the slope parameters β in model M_j conditional on other parameters is given by

$$p(\beta|\sigma^2, \mathbf{\Omega}, M_j) \sim N(\beta_j, \sigma_j^2 \mathbf{V}_j) \quad (19)$$

where β_j and \mathbf{V}_j , denote, respectively, the conditional posterior mean and variance. The conditional mean of β is estimated using the generalized least squares (GLS) estimator

$$\beta_j = \mathbf{V}_j \left(\mathbf{X}'_j \mathbf{\Omega}_j^{-1} \mathbf{y} \right). \quad (20)$$

Intuitively, the robust GLS estimator weights each observation of the dependent variable y_i and regressors $\mathbf{X}_{i,j}$ by the inverse of their estimated error variance $\omega_{i,j}$, which is the diagonal element of the variance matrix $\mathbf{\Omega}_j$. This is in contrast to estimation by ordinary least squares (OLS) in the benchmark normal case, where each observation receives the same weight.

The posterior error variances are estimated as follows. The posterior distribution of the common error variance parameter σ_j^2 conditional on other parameters is calculated using the weighted sum of squared errors, with weights proportional to the error variance $\omega_{i,j}$ for each observation $i = 1, \dots, N$. The degrees of freedom is equal to N and not $N - k_j$, since we condition on the slope parameters β_j and the error covariance matrix $\mathbf{\Omega}_j$. The elements of the error variance matrix $\mathbf{\Omega}_j$ conditional on other parameters are drawn from their posterior Chi-square distribution, where the $v + 1$ degrees of freedom follow from combining the prior distribution of $\omega_{i,j}$ (17) with v degrees of freedom with the weighted sum of squared error kernel from the likelihood function with one degree of freedom, conditional on β_j and σ_j^2 .

The degrees of freedom v governing the weight on outliers are determined by the following alternatives. Under case (a), the degrees of freedom are fixed to their constant prior values \bar{v} . Under case (b), we use the exponential prior distribution (18) with prior mean $v_0 = 25$, and draw degrees of freedom v from the posterior distribution. Since the posterior objects of interest are much less affected by the hierarchical prior for the degrees of freedom in case (b), this is our preferred specification for robust model averaging. In particular, the posterior results are not sensitive to the choice of the parameter v_0 , so we have chosen this parameter to allow for moderate heteroscedasticity a priori.

The estimation of the posterior objects of interest that are unconditional with respect to the space of models is analogous to the benchmark model averaging case in section 2.2, except that we now use robust (GLS) estimates of conditional means and variances in each model M_j . In contrast to a simple correction of standard errors, *all* posterior objects of interest are affected by heterogeneity and outliers through the different weights on each observation. Conditional on inclusion, posterior means and standard deviations are calculated by (13) and (14), respectively. The posterior inclusion probability associated with each variable is calculated by summing the posterior model probability (10) when the variable is included. We still use the posterior (BIC) model weights (8), which are approximations to Bayesian posterior weights. Section 4.3 also considers proper Bayesian weights as alternative.

4 Robustness of Growth Determinants

This section investigates the robustness of statistical inference and economic analysis of growth determinants when simultaneously confronted with model uncertainty, parameter heterogeneity and outliers. Results obtained using benchmark model averaging are contrasted with findings based on robust model averaging. First, the data and benchmark model averaging results of Sala-i-Martin et al (2004) are summarized. Second, robust model averaging is applied to the same data on growth determinants. This implies interesting differences in the posterior distributions associated with the associated slope coefficients, measuring the economic effect of growth determinants. In particular, changes in posterior inclusion probabilities and inference are investigated. The posterior distribution of error variances for each observations is estimated, which allows the identification of important outliers in the cross-country growth data. Third, the section presents sensitivity analysis with respect to prior assumptions.

4.1 Data and Benchmark Results

The empirical growth literature has proposed a large number of explanatory variables related to economic growth (see for example the variables listed in Durlauf and Quah, 1999; or Durlauf *et al.*, 2008). To facilitate comparison with results obtained by benchmark model averaging described in section 2, we apply the robust model averaging proposed in section 3 to the dataset by Sala-i-Martin *et al.* (2004). Sala-i-Martin *et al.* (2004) select variables representing ‘state variables’ in economic growth models and measure them as close as possible to the start of the sample period in 1960. Furthermore, the dataset is restricted to be balanced, i.e. without missing observations. Under these criteria the total number of explanatory variables is $K = 67$ with observations for $N = 88$ countries. The dependent variable, *average growth rate of GDP per capita between 1960-96*, and the 67 explanatory variables are listed in the Data Appendix B. The Data Appendix also lists short names of variables, brief descriptions of variables, and sample mean and standard deviations.¹³

For comparison with Sala-i-Martin *et al.* (2004), explanatory variables in the Data Appendix B and in all Tables with results are ordered by their *Posterior Inclusion Probability* (PIP) in the benchmark normal case. Recall that the posterior inclusion probability, defined in equation (10), is an overall measure of the importance of explanatory variables in explaining economic growth. The posterior inclusion probability can be compared to the prior inclusion probability, which equals $\pi_k = \bar{k}/K = 7/67 = 0.104$ for prior model size $\bar{k} = 7$. For comparison with benchmark results by Sala-i-Martin *et al.*, all results discussed in this section assume a *fixed* prior model size. Sensitivity analysis using random prior model size are discussed in section 4.3.

Using the benchmark model averaging approach from section 2, Sala-i-Martin *et al.* (2004) call the 18 highest ranked explanatory variables with posterior inclusion probability greater than the prior probability, ‘significantly’ related to economic growth.

¹³The dataset and complete tables of results are available at: www.nhh.no/sam/bace.

The next three variables, ranked 19 to 21, are found to be ‘marginally’ partially related to economic growth. The following discussion contrasts the benchmark results by Sala-i-Martin *et al.* with those using robust model averaging, with particular attention to the 21 highest ranked variables.

4.2 Main Robust Findings

Table 1 contrasts the the posterior inclusion probabilities for the benchmark normal benchmark case, shown in column (1), with estimates using robust model averaging, shown in column (2). For most explanatory variables, posterior inclusion probabilities are smaller when estimated robustly, indicating that a researchers confidence about the inclusion of regressors is affected by heteroscedastic errors due to heterogeneity and outliers. For convenience, variables with posterior inclusion probability below the prior cut-off of $\pi_k = 0.104$ are highlighted in *italics* in Table 1. The posterior mean degrees of freedom $E(v|\mathbf{y})$ in the robust model averaging case with random degrees of freedom equal 19.5 – compared with a prior value of 25 –, implying important deviations from the homoscedastic normal benchmark case and evidence for the presence of outliers. Provided that a researcher has strong priors on the likely presence of outliers, Table 1 also shows posterior inclusion probabilities for *fixed* prior degrees of freedom ($\bar{v} = 100, 25, 5$), with increasing *prior* weight on outliers as we move from column (3) to column (5).

[INSERT TABLE 1 ABOUT HERE]

Figures 2 to 7 contrasts the posterior distributions for the 24 highest-ranked variables under the normal benchmark case with the posterior distributions estimated using robust model averaging. Notice that the posterior distributions consist of two parts: The first part is a mass at zero (the green bars) measure the posterior probability that the variable in question is not included in the regression model.¹⁴ The second part shows the posterior distributions using either normal benchmark model averaging (blue), or alternatively using robust model averaging (red). For most growth determinants, the robust posterior distributions are scaled down relative to the benchmark normal case, indicating that the posterior probability of inclusion is lowered for many variables. Most posterior distributions of slope coefficients also shrink towards zero for most variables when using robust model averaging, indicating smaller average coefficients (in absolute value) and economic effects when estimated robustly.

[INSERT FIGURES 2–7 ABOUT HERE]

Table 2 shows the posterior mean and standard deviation of slope coefficients of growth determinants. For comparison with Classical inference, the Table shows estimates *conditional on including* variable \mathbf{x}_k in the regression model.¹⁵ Column (1) shows

¹⁴Note that for ease of presentation, we have split the point mass at zero into seven bars. Also the bars are truncated above for some lower ranked variables that have low posterior inclusion probability.

¹⁵The unconditional moments that reflect also uncertainty about inclusion of a variable, can readily be calculated from the conditional ones, as discussed in section 2.2.

the posterior mean and underneath the standard deviation of the slope coefficient for the benchmark model averaging case. In contrast, column (2) shows the posterior estimates using robust model averaging. Comparing these two columns, we can observe that for most explanatory variables the ratio of posterior mean to standard deviation is smaller in absolute value for most regressors under robust model averaging compared to the benchmark case. An absolute value much below two of such standardized coefficients indicates that the coefficient is not considered to be statistically significant according to Classical hypothesis tests (see also Brock and Durlauf, 2001). For convenience, cases where the ratio of posterior mean to standard deviation falls below two are highlighted by setting them in *italics*. To compare economic inference with the situation where there exists strong priors on the presence of outliers, Table 2 also show the posterior mean and standard deviation conditional on inclusion of each regressor for *fixed* prior degrees of freedom ($\bar{v} = 100, 25, 5$), with increasing *prior* weight on outliers as we move from column (3) to column (5).

[INSERT TABLE 2 ABOUT HERE]

The robust model averaging procedure allows the explicit investigation of outliers in the economic growth dataset. Figure 8 shows the diagonal entries of the posterior error variance matrix $E(\Omega|\mathbf{y}) = E(diag(\omega_i)|\mathbf{y})$. In this dataset, Botswana has a posterior variance more than three times larger than the average over the 67 other country in this sample. Observe that Botswana has an average annual growth rate of 4.7 percent during the 1960-96 sample period, which is much higher than neighboring countries in Sub-Saharan Africa. Other outlying observations are the Philippines with posterior variance about twice the average, and also the Central African Republic, Gabon, Zaire and Zambia. It is noticeable that five out of six countries with strong outliers are located in Africa, and in particular Sub-Saharan Africa. Notice that these outliers are present *despite* the inclusion of the Sub-Saharan Africa dummy as regressor in many models.

[INSERT FIGURE 8 ABOUT HERE]

The main findings regarding the robustness of particular growth determinants can be summarized as follows.

Robust Growth Determinants – The eight highest ranked regressors in Table 1 have posterior inclusion probabilities exceeding the prior probability of 0.104 under all the different prior assumptions. Among these robust variables, are the *East Asian* dummy and variables measuring natural conditions, such as the *Fraction of Tropical Area* and *Malaria Prevalence*. Other variables can be viewed as state variables in standard neoclassical growth models, such as *Primary Schooling Enrolment*, the *Price of Investment Goods*, and *Initial Income*. Two additional robust variables measure initial conditions, such as the *Coastal Population Density* and *Life Expectancy*. Regarding statistical inference and economic significance, Table 2 shows that for these eight robust growth determinants, posterior mean and standard deviation conditional on inclusion are very similar under the benchmark and robust model averaging. For example, the

robust estimate for the speed of convergence of one percent implied by the estimated mean coefficient -0.0082 is close to the benchmark case. A researcher can therefore conclude that the eight highest ranked explanatory variables are robust determinants of economic growth, even when allowing for the presence of sizeable outliers *a priori*, parameter heterogeneity and model uncertainty.

Growth Determinants Sensitive to Heteroscedasticity – The other regressors up to rank 18 in Table 1 are sensitive to some deviations from the normal benchmark case, when we allow for the presence of outliers and heterogeneity. Statistical and economic inference about the relevance of these regressors are also affected by allowing for heteroscedastic errors as shown by the italicized entries in Table 2. The population *Fraction Confucian* and the dummy variable for former *Spanish Colony* – have posterior inclusion probability higher than the prior probability in most cases in Table 1. Also ratio of posterior mean to standard deviation of associated coefficients shown in Table 2 are above or close to two, indicating confidence that these two regressors are important in explaining economic growth. Inference about dummy variables for *Sub-Saharan African* and *Latin American* is affected by allowing for different degrees of heteroscedasticity. This is perhaps not too surprising, because the dummy variables might to some extent capture outlying country observations. Similarly, the *Fraction Muslim* and *Fraction Buddhist* capture some of the heterogeneity of growth performance that is partly soaked up by allowing for heteroscedastic errors. The *Fraction GDP in Mining* is strongly affected by allowing for heteroscedastic errors, since its effect is not considered significant under *any* deviation from the normal benchmark model averaging case. Observe that Botswana is indentified as important outlier in the sample. Botswana also has a large mining share of GDP, so the coefficient on the Mining share might have picked up the outlier in the benchmark case.¹⁶ Three explanatory variables, the *Number of Years an Economy has Been Open*, *Ethnolinguistic Fractionalization*, and the *Government Consumption Share* of GDP have posterior inclusion probability already close to the prior cutoff value of 0.104, and ratio of absolute value of posterior mean to standard deviation near or below 2 in the normal benchmark case.

Regressors Marginally or Unrelated with Growth – Finally, three marginal variables (ranked 19 to 21), *Population Density*, *Real Exchange Rate Distortions* and *Fraction Speaking a Foreign Language* are marginally related to economic growth under the normal benchmark case. Estimating their effect robustly further weakens confidence in their importance. We confirm the finding of Sala-i-Martin *et al.* (2004) that none of the variables ranked 22-67 are found to be related to economic growth regardless of whether benchmark and robust version of model averaging are being used.

4.3 Sensitivity Analysis

In order to determine the impact of a number of critical prior assumptions, we examine the sensitivity of our results to following deviations from the prior structure in the

¹⁶We thank Andrew Warner for pointing out this fact.

previous section.

First, we compare the benchmark results based upon BIC model weights (8) with those using proper Bayesian conjugate priors (see the Computational Appendix A for implied posterior model weights). To contrast the results with the normal benchmark case that uses sample-dominated priors, we set the degrees of freedom parameters to a relatively large value ($\bar{v} = 100$), which implies that the error distribution is approximately normal. This implies a thin-tailed distribution for the errors with low prior weight on outliers. The results using posterior (Bayes) estimates are shown in column (8) in Tables 1 and 2. The classification of explanatory variables is very similar to the robust model averaging: regressors ranked 1 to 8 are robustly related with posterior inclusion probability exceeding the prior probability, whereas regressors ranked below 9 have lower probability of inclusion after averaging across models. Similar conclusions can be drawn about statistical and economic inference about significance of effects in column (9) of Table 2. An exception is the *Spanish Colony* dummy, which shows an *increase* in posterior inclusion probability and absolute value of standardized coefficients when comparing columns (1) and (8) in Tables 1 and 2, respectively. The conclusions reached using proper Bayesian priors are therefore in line with the ones using robust model averaging.

Second, we also contrast the standard errors estimated under the normal benchmark case with the heteroscedasticity-consistent standard errors (HCSE), proposed by White (1980). For each model M_j , the covariance matrix for the slope coefficient β_k can be consistently estimated by

$$V(\beta_k|\mathbf{y}, M_j) = (\mathbf{X}'_j \mathbf{X}_j)^{-1} \mathbf{X}'_j \boldsymbol{\Omega}_j \mathbf{X}_j (\mathbf{X}'_j \mathbf{X}_j)^{-1} \quad (21)$$

This estimator is unbiased asymptotically, but has been shown to be biased in small samples. We therefore use the simple finite-sample correction of scaling the covariance estimator $\hat{\Omega}_j = SSE_j \cdot N/(N - k_j)$ (see MacKinnon and White 1985). Column (9) in Tables 2 and 4 shows the posterior standard errors when using White's robust standard errors (21) and benchmark normal posterior model weights. Note that the coefficient means and therefore posterior model weights are not affected by estimating standard errors using HCSE. Interestingly, the standardized coefficients in column (9) based on the White heteroscedasticity-corrected standard errors do not seem to reflect the non-robustness of several growth determinants highlighted by robust model averaging using either random or fixed degrees of freedom. On the contrary, the sole use of White standard errors in equation (21) could give a misleading indication of robustness, when contrasting inference based on standardized coefficients in column (9) of Table 2 with posterior inclusion probabilities in column (3) of Table 1. We therefore recommend to use estimates based on HCSE with caution.

Third, we investigate the sensitivity of our results to relaxing the assumption of a fixed prior model size. As discussed in section 2.1, we use a hierarchical (Beta-Binomial) prior for the prior model size parameter \bar{k} . The results from Tables 1 and 2 are reproduced for random model sizes in Tables 3 and 4. Comparing the results under

fixed and random prior model size, we can draw the following conclusions:

- The robustly related variables ranked 1-8 have posterior inclusion probabilities above the prior under all variations of prior assumption. We can also conclude that inference on statistical and economic significance is robust to using different estimates of posterior means and standard deviations.
- For the sensitive variables, a similar qualitative picture emerges, namely robustly estimated PIP's are lower than under normal benchmark assumptions, and standardized coefficients smaller in absolute value. Notice however, that PIPs are higher using random model size.
- For the marginally and not related variables we draw similar conclusions under alternative prior assumptions.

5 Conclusion

This paper investigates the sensitivity of the benchmark (Bayesian) Model Averaging (BMA) procedure. We evaluate the robustness of the benchmark model averaging in the context of cross-country growth regressions and find that inference on the robustness of growth determinants is significantly affected by considering deviations from the benchmark assumptions, such as outliers and heteroscedastic errors.

In contrast to the results based on benchmark model averaging, we find that statistical inference and the economic importance of a number of variables is sensitive to allowing for heteroscedasticity due to outliers or parameter heterogeneity a priori.

We are working on extending the robust model averaging approach to a broader set of prior parameters and to compare our approach in more detail with alternative methods.

References

- [1] Brock W, Durlauf S. 2001. Growth Empirics and Reality. *World Bank Economic Review* 15(2): 229-72.
- [2] Brown PJ, Vannucci M, Fearn T. 1998. Multivariate Bayesian Variable Selection and Prediction. *Journal of the Royal Statistical Society B*, 60(3): 627-41.
- [3] Canova F. 2004. Testing for Convergence Clubs in Income per Capita: A Predictive Density Approach. *International Economic Review* 45(1): 49-77.
- [4] Chib S. 1993. Markov Chain Monte Carlo Methods: Computation and Inference. In *Handbook of Econometrics*, Vol. 5, Heckman JJ, Leamer EE. (eds.), Elsevier: North Holland.
- [5] Ciccone A, Jarocinski M. 2010. Determinants of Economic Growth: Will Data Tell? *American Economic Journal: Macroeconomics*, Forthcoming.
- [6] Clyde M.A, Lee H. 2001. Bagging and the Bayesian bootstrap. In *Artificial Intelligence and Statistics*, 169174, Richardson T, Jaakkola T. (eds.), Proceedings of the Eighth International Workshop, Morgan Kaufmann Publishers.
- [7] Corrado L, Martin R, Weeks M. 2005. Identifying and Interpreting regional Convergence Clusters Across Europe. *Economic Journal* 115: C133-C160.
- [8] Crespo Cuaresma J, Doppelhofer G, Feldkircher M. 2009. The determinants of Economic Growth in European Regions. *CESifo Working Paper* No. 2519.
- [9] Deaton A. 2010. Price Indexes, Inequality, and the Measurement of World Poverty. *American Economic Review* 100(1): 5-34.
- [10] Doppelhofer G. 2008. Model Averaging. *The New Palgrave Dictionary of Economics*. Second Edition. Durlauf SN, Blume LE. (eds.), Palgrave Macmillan. <http://www.dictionaryofeconomics.com/>
- [11] Doppelhofer G, Weeks M. 2009. Jointness of Growth Determinants. *Journal of Applied Econometrics* 24: 209-44.
- [12] Durlauf SN, Johnson P. 1995. Multiple Regimes and Cross Country Growth Behaviour. *Journal of Applied Econometrics* 10(4): 365-84.
- [13] Durlauf SN, Johnson PA, Temple JRW. 2008. Growth Econometrics. Aghion P and Durlauf SN (eds.) *Handbook of Economic Growth*. Elsevier: North Holland.
- [14] Eicher TS, Papageorgiou C, Raftery AE. 2009. Default Priors and Predictive Performance in Bayesian Model Averaging, with Applications to Growth Determinants. *Journal of Applied Econometrics*, Forthcoming.

- [15] Feldkircher M, Zeugner S. 2009. Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging. *IMF Working Paper* WP/09/202.
- [16] Fernandez C, Ley E, Steel MFJ. 2001a. Model Uncertainty in Cross-Country Growth Regression. *Journal of Applied Econometrics* 16: 563-76.
- [17] Fernandez C, Ley E, Steel MFJ. 2001b. Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics* 100(2): 381-427.
- [18] Fernandez C, Steel MFJ. 2000. Bayesian Regression Analysis with Scale Mixtures of Normals. *Econometric Theory* 16: 80-101.
- [19] George E, McCulloch R. 1993. Variable Selection Via Gibbs Sampling. *Journal of the American Statistical Association* 88(423): 881-89.
- [20] Geweke J. 1993. "Bayesian Treatment of the Independent Student- t Linear Model." *Journal of Applied Econometrics* 8: S19-S40.
- [21] Gottardo R, Raftery AE. 2007. Bayesian Robust Variable and Transformation Selection: A Unified Approach. *Technical Report* no. 508. Department of Statistics. University of Washington.
- [22] Hendry DF, Santos C. 2005. Regression Models with Data-Based Indicator Variables. *Oxford Bulletin of Economics and Statistics* 67(5): 571-95.
- [23] Hoeting J, Raftery A, Madigan D. 1996. A Method for Simultaneous Variable Selection and Outlier Identification in Linear Regression. *Computational Statistics and Data Analysis* 22: 251-70
- [24] Hoeting J, Madigan D, Raftery A, Volinsky C. 1999. Bayesian Model Averaging: A Tutorial. *Statistical Science* 14: 382-417.
- [25] Jeffreys H. 1961. *Theory of Probability*. Third edition. Oxford University Press: London.
- [26] Kass R, Raftery A. 1995. Bayes Factors. *Journal of the American Statistical Association* 90(430): 773-95.
- [27] Kass R, Wasserman L. 1995. A Reference Bayesian Test for Nested Hypotheses and its Relationship to the Schwarz Criterion. *Journal of the American Statistical Association* 90: 928-34.
- [28] Klein RW, Brown SJ. 1984. Model Selection When There is "Minimal" Prior Information. *Econometrica* 52(5): 1291-321.
- [29] Koop G. 2003. *Bayesian Econometrics*. Wiley: Chichester.

- [30] Lange KL, Little RJA, Taylor JMG. 1989. Robust Statistical Modeling Using the t Distribution. *Journal of the American Statistical Association* 84(408): 881-96.
- [31] Leamer EE. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. Wiley: New York.
- [32] Leamer E.E. 1982. Sets of Posterior Means with Bounded Variance Priors. *Econometrica* 50(3): 725-36.
- [33] Leamer EE, Leonard H. 1983. Reporting the Fragility of Regression Estimates. *The Review of Economics and Statistics* 65: 306-17.
- [34] LeSage JP. and Parent, O. 2007. Bayesian Model Averaging for Spatial Econometric Models. *Geographical Analysis* 39(3): 241-67.
- [35] Levine RE, Renelt D. 1992. A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review* 82(4): 942-63.
- [36] Ley E, Steel MFJ. 2009. On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regressions. *Journal of Applied Econometrics* 24: 651-74.
- [37] MacKinnon J.G, White H. 1985. Some Heteroscedasticity- Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. *Journal of Econometrics* 29: 305-25.
- [38] Magnus JR, Wan TK and Zhang X. 2010. WALS estimation with Nonspherical Disturbances and an Application to the Hong Kong Housing Market. mimeo.
- [39] Masanjala W, Papageorgiou C. 2008. Rough and Lonely Road to Prosperity: A Re-examination of the Sources of Growth in Africa Using Bayesian Model Averaging. *Journal of Applied Econometrics* 23: 671-82.
- [40] McLachlan G, Peel D. 2000. *Finite Mixture Models*. Wiley: New York.
- [41] Mitchell T, Beauchamp J. 1988. Bayesian Variable Selection in Linear Models. *Journal of the American Statistical Association* 83(404): 1023-32.
- [42] Papageorgiou C. (ed.) 2007. Special Issue on the Empirics of Growth Nonlinearities, *Journal of Macroeconomics* 29(3).
- [43] Raftery A. 1995. Bayesian Model Selection in Social Research. *Sociological Methodology* 25, 111-63.
- [44] Rousseeuw PJ. 1984. Least Median of Squares Regression. *Journal of the American Statistical Association* 79: 871-80.
- [45] Rubin DB. 1981. The Bayesian Bootstrap. *Annals of Statistics* 9: 130-34.

- [46] Sala-i-Martin X, Doppelhofer G, Miller RI. 2004. Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review* 94(4): 813-35.
- [47] Schultz TP. 1999. Health and Schooling Investments in Africa. *Journal of Economic Perspectives* 13(3): 67-88.
- [48] Schwarz G. 1978. Estimating the Dimensions of a Model.” *Annals of Statistics* 6(2): 461-4.
- [49] Sturm JE, de Haan J. 2005. Determinants of Long-Term Growth: New results Applying Robust Estimation and Extreme Bounds. *Empirical Economics* 30: 597-617.
- [50] Tan CM. 2009. No One True Path: Uncovering the Interplay between Geography, Institutions, and Fractionalization in Economic Development. *Journal of Applied Econometrics*. Forthcoming.
- [51] Temple J. 2000. Growth Regressions and what the Textbooks Don’t Tell You. *Bulletin of Economic Research* 52(3): 181-205.
- [52] Wasserman L. 2000. “Bayesian Model Selection and Model Averaging.” *Journal of Mathematical Psychology* 44: 92-107.
- [53] White H. 1980. A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica* 48(4): 817-38.
- [54] Zaman A, Rousseeuw P.J. and Orhan M. 2001. Econometric Applications of High-Breakdown Robust Regression Techniques. *Economics Letters* 71: 1-8.
- [55] 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, Zellner A. (eds.). North-Holland: Amsterdam.

A Computational Appendix

This Appendix discusses some of the details of the sampling and model averaging procedure: (1) the random and stratified sampling procedures across the model space, (2) the Gibbs sampler used in the robust model averaging discussed in section 3.1, and (3) numerical convergence criteria to check convergence of the sampler.

A.1 Random and Stratified Sampling

For the benchmark normal case of section 2 we estimate the posterior mean and variance of β_j conditional on each model M_j by OLS. For the sensitivity checks in section 4.3, we can also calculate analytically the marginal likelihood (32).

The *random sampler* therefore draws directly from the posterior distribution of $p(\beta|\mathbf{y})$. With $K = 67$ regressors the number of possible regression models equals $2^{67} \approx 1.48 \times 10^{20}$. Each regression takes approximately 0.0005 seconds using GAUSS on a recent PC. An exhaustive search over all models is therefore not feasible. Instead we draw from the posterior distribution and related parameters of interest until the parameters have converged (see numerical convergence criteria below).

The *stratified sampler* introduced by Sala-i-Martin *et al.* (2004, Technical Appendix) allows for sampling inclusion probabilities to differ from prior inclusion probabilities $\pi = \bar{k}/K$. After every 100,000 regressions, the sampling inclusion probability for each variable π_i^S are set equal to a weighted average (with weight 0.5 in the benchmark case) of the initial inclusion probability and the posterior inclusion probability (4) estimated on those runs. To avoid sampling only a very small set of variables, the sampling inclusion probabilities are restricted to lie in the interval $[0.1, 0.85]$. The stratified sampler over-samples models that include variables with high inclusion probability which greatly speeds up numerical convergence. To correct for differences in sampling probabilities of different regressors, we scale the posterior weights by the ratio of prior to sampling probabilities π/π_i^S for $i = 1, \dots, K$.

A.2 Gibbs Sampling

The robust approach adopted in section 3 implies the following hierarchical structure, which can be conveniently estimated by the Gibbs sampler (see also Geweke, 1993).

1. The posterior distribution of the slope parameters β in model M_j conditional on other parameters is given by

$$p(\beta|\sigma^2, \mathbf{\Omega}, M_j) \sim N(\beta_j, \sigma_j^2 \mathbf{V}_j). \quad (22)$$

The conditional mean of β_j is estimated using the generalized least squares (GLS) estimator

$$\beta_j = \mathbf{V}_j \left(\mathbf{X}_j' \mathbf{\Omega}_j^{-1} \mathbf{y} \right). \quad (23)$$

The posterior variance of β_j is given by $\sigma_j^2 \bar{\mathbf{V}}_j$, where

$$\bar{\mathbf{V}}_j = \left(\mathbf{X}'_j \boldsymbol{\Omega}^{-1} \mathbf{X}_j + \sigma_j^2 \mathbf{V}_{0j}^{-1} \right)^{-1}. \quad (24)$$

combines both sample and prior information given by the g -prior for the variance \mathbf{V}_{0j} .

2. The posterior distribution of the error variance parameter σ_j^2 conditional on the other parameters is given by

$$\left[\sum_{i=1}^N (e_{i,j}^2 / \omega_{i,j}) \right] / \sigma_j^2 \mid (\beta_j, \boldsymbol{\Omega}_j, M_j) \sim \chi^2(N), \quad (25)$$

where $e_{i,j} = y_i - \mathbf{x}'_{i,j} \beta_j$. Notice that the degrees of freedom equals N and not $N - k_j$, since we condition on β_j .

3. The posterior distribution of the elements of the error variance matrix $\boldsymbol{\Omega}_j$ conditional on the other parameters is proportional to

$$\left[\left(\sigma_j^{-2} e_{i,j}^2 + v \right) / \omega_{i,j} \right] \mid (\beta_j, \sigma_j^2, M_j) \sim \chi^2(v+1), \quad i = 1, \dots, N. \quad (26)$$

The $v+1$ degrees of freedom follow from combining the prior distribution of $\omega_{i,j}$ (17) with v degrees of freedom with terms $\sigma^{-2} e_{i,j}^2 / \omega_{i,j}$ from the likelihood function with a $\chi^2(1)$ kernel.

4. For the degrees of freedom parameter v we consider two choices:

- (a) Either we set the degrees of freedom to fixed values $\bar{v} = 5, 25, 100$.
- (b) Alternatively, we draw degrees of freedom v from its posterior distribution:

$$p(v \mid \omega_{i,j}, M_j) \sim p(\omega_{i,j} \mid v, M_j) \cdot p(v) \quad (27)$$

$$\propto \left(\frac{v}{2} \right)^{Nv/2} \cdot \Gamma(v/2)^{-N} \cdot \exp(-\eta v) \quad (28)$$

where $\eta \equiv \sum_{i=1}^N [\ln \omega_{i,j} + 1/\omega_{i,j}] / 2 + 1/v_0$. This distribution has no convenient analytic form, and we therefore introduce a Metropolis step to draw candidate values of v' from (28) and accept them with probability $p_\alpha = 0.5$. For an introductory discussion of the Metropolis-Hastings algorithm, see Chib (2001).

For the robust case of section 3.1, analytic expressions of the marginal likelihood are not available. However, we can make a sufficiently large number of draws from the *conditional* posterior distributions (22)-(24) for mean and variance of the slope coefficient β_j , (25) and (26) for the error variance $\sigma_j^2 \boldsymbol{\Omega}_j$, and from the posterior distribution (28) for the degrees of freedom v . Alternatively, we set the degrees of freedom to fixed values \bar{v} .

A.3 Numerical Convergence Criteria

Under certain regularity conditions the chain of draws from the Gibbs sampler converges to the posterior distribution as the number of draws becomes large (see Chib, 2001). Let $s = 1, \dots, S$ be number of draws from the posterior distributions. We discard S_0 burn-in draws and estimate the parameters using the remaining $S_1 = S - S_0$ draws.

- The posterior mean $\bar{\beta}_j \equiv E(\beta_j | M_j, \mathbf{y})$ is estimated by $\frac{1}{S_1} \sum_{s=S_0+1}^S \beta_{j,s}$, using draws from (22).
- The posterior variance matrix $\bar{\mathbf{V}}_j \equiv Var(\beta_j | M_j, \mathbf{y})$ can be calculated either numerically (superscript n) or analytically (superscript a) using

$$\bar{\mathbf{V}}_j^n = \frac{1}{S_1} \sum_{s=S_0+1}^S \beta_{j,s}^2 - [E(\beta_j | M_j, \mathbf{y})]^2 \quad (29)$$

$$\bar{\mathbf{V}}_j^a = \left(\mathbf{X}_j' \boldsymbol{\Omega}_j^{-1} \mathbf{X}_j + \sigma_j^2 \mathbf{V}_{0j}^{-1} \right)^{-1} \quad (30)$$

where $\bar{\sigma}_j^2 \equiv E(\sigma^2 | M_j, \mathbf{y})$ is estimated by $\frac{1}{S_1} \sum_{s=S_0+1}^S \sigma_{j,s}^2$ using draws from (25).

- The posterior error variance $\bar{\boldsymbol{\Omega}}_j \equiv E(\boldsymbol{\Omega} | M_j, \mathbf{y})$ is estimated using draws from (26)

$$\frac{1}{S_1} \sum_{s=S_0+1}^S \boldsymbol{\Omega}_{j,s} = \frac{1}{S_1} \sum_{s=S_0+1}^S \text{diag}(\omega_{1,j,s}, \dots, \omega_{N,j,s}) \quad (31)$$

We combine the robust estimates from the Gibbs sampler for the conditional mean, $\bar{\beta}_j$, and conditional variance, $\bar{\sigma}_j^2 \bar{\mathbf{V}}_j$, with posterior (BIC) model weights (7) to estimate the posterior mean and variance unconditionally over the model space \mathcal{M} . For the sensitivity analysis in section 4.3, we also consider proper Bayesian normal-inverse Gamma prior, implying that the marginal likelihood of model M_j is proportional to

$$l(\mathbf{y} | M_j) \propto \left(\frac{|\bar{\mathbf{V}}_j|}{|\mathbf{V}_{0j}|} \right)^{1/2} \cdot (N \cdot \sigma_j^2)^{-N/2} \quad (32)$$

To check for *numerical convergence* of model averaging with Gibbs sampling, we observe that for a large enough number of draws S from the Gibbs sampler, the numerical (29) and analytic (30) estimates of the posterior variance should be similar. The default values for robust model averaging are 200 draws from the Gibbs sampler (with 20 draws for burnin) and 10 million draws from the model averaging loops. We use the difference between posterior standardized coefficients using the numerical and analytic posterior variance as numerical convergence criterion in this case. In particular, we choose the number of Gibbs draws \bar{S} , so that

$$|\beta_k / \sqrt{\mathbf{V}_k^n} - \beta_k / \sqrt{\mathbf{V}_k^a}|_{S=\bar{S}} < 0.1, \quad k = 1, \dots, K \quad (33)$$

B Data Appendix

Rank	Short Name	Variable Description	PIP	Mean	S.D.
Depend. Variable	GROWTH	Average Growth Rate of PPP-adjusted GDP per Capita between 1960–1996	–	0.0182	0.019
1	EAST	East Asian Dummy	0.823	0.1136	0.3192
2	P60	Primary Schooling Enrollment	0.796	0.7261	0.2932
3	IPRICE1	Investment Price	0.774	92.47	53.68
4	GDPCH60L	Initial income (Log GDP in 1960)	0.685	7.3549	0.9011
5	TROPICAR	Fraction of Tropical Area	0.563	0.5702	0.4716
6	DENS65C	Population Coastal Density	0.428	146.87	509.83
7	MALFAL66	Malaria Prevalence	0.252	0.3394	0.4309
8	LIFE060	Life Expectancy	0.209	53.72	12.06
9	CONFUC	Fraction Confucian	0.206	0.0156	0.0793
10	SAFRICA	Sub-Saharan Africa Dummy	0.154	0.3068	0.4638
11	LAAM	Latin American Dummy	0.149	0.2273	0.4215
12	MINING	Fraction GDP in Mining	0.124	0.0507	0.0769
13	SPAIN	Spanish Colony Dummy	0.123	0.1705	0.3782
14	YRSOPEN	Years Open 1950-94	0.119	0.3555	0.3444
15	MUSLIM00	Fraction Muslim	0.114	0.1494	0.2962
16	BUDDHA	Fraction Buddhist	0.108	0.0466	0.1676
17	AVELF	Ethnolinguistic Fractionalization	0.105	0.3476	0.3016
18	GVR61	Government Consumption Share	0.104	0.1161	0.0745
19	DENS60	Population Density	0.086	108.07	201.44
20	RERD	Real Exchange Rate Distortions	0.082	125.03	41.71
21	OTHFRAC	Fraction Speaking Foreign Language	0.080	0.3209	0.4136
22	OPENDEC1	Openness 1965-74	0.076	0.5231	0.3359
23	PRIGHTS	Political Rights	0.066	3.8225	1.9966
24	GOVSH61	Government Share of GDP	0.063	0.1664	0.0712
25	H60	Higher Education Enrollment	0.061	0.0376	0.0501
26	TROPPOP	Fraction Population In Tropics	0.058	0.3000	0.3731
27	PRIEXP70	Primary Exports	0.053	0.7199	0.2827
28	GGCFD3	Public Investment Share	0.048	0.0522	0.0388
29	PROT00	Fraction Protestant	0.046	0.1354	0.2851
30	HINDU00	Fraction Hindu	0.045	0.0279	0.1246
31	POP1560	Fraction Population Less than 15	0.041	0.3925	0.0749
32	AIRDIST	Air Distance to Big Cities	0.039	4324	2614
33	GOVNOM1	Nominal Government Share	0.036	0.1490	0.0584
34	ABSLATIT	Absolute Latitude	0.033	23.21	16.84
35	CATH00	Fraction Catholic	0.033	0.3283	0.4146
36	FERTLDC1	Fertility	0.031	1.5620	0.4193
37	EUROPE	European Dummy	0.030	0.2159	0.4138
38	SCOUT	Outward Orientation	0.030	0.3977	0.4922
39	COLONY	Colony Dummy	0.029	0.7500	0.4355
40	CIV72	Civil Liberties	0.029	0.5095	0.3259
41	REVCoup	Revolutions and Coups	0.029	0.1849	0.2322
42	BRIT	British Colony Dummy	0.027	0.3182	0.4684
43	LHCPC	Hydrocarbon Deposits	0.025	0.4212	4.3512
44	POP6560	Fraction Population Over 65	0.022	0.0488	0.0290
45	GDE1	Defense Spending Share	0.021	0.0259	0.0246
46	POP60	Population in 1960	0.021	20308	52538
47	TOT1DEC1	Terms of Trade Growth in 1960s	0.021	-0.0021	0.0345
48	GEEREC1	Public Education Spending Share	0.021	0.0244	0.0096
49	LANDLOCK	Landlocked Country Dummy	0.021	0.1705	0.3782
50	HERF00	Religion Measure	0.020	0.7803	0.1932
51	SIZE60	Size of Economy	0.020	16.15	1.82
52	SOCIALIST	Socialist Dummy	0.020	0.0682	0.2535
53	ENGFRAC	English Speaking Population	0.020	0.0840	0.2522
54	PI6090	Average Inflation 1960-90	0.020	13.13	14.99
55	OIL	Oil Producing Country Dummy	0.019	0.0568	0.2328
56	DPOP6090	Population Growth Rate 1960-90	0.019	0.0215	0.0095
57	NEWSTATE	Timing of Independence	0.019	1.0114	0.9767
58	LT100CR	Land Area Near Navigable Water	0.019	0.4722	0.3802
59	SQPI6090	Square of Inflation 1960-90	0.018	394.54	1119.70
60	WARTIME	Fraction Spent in War 1960-90	0.016	0.0695	0.1524
61	LANDAREA	Land Area	0.016	867189	1814688
62	ZTROPICS	Tropical Climate Zone	0.016	0.1900	0.2687
63	TOTIND	Terms of Trade Ranking	0.016	0.2813	0.1904
64	ECORG	Capitalism	0.015	3.4659	1.3809
65	ORTH00	Fraction Orthodox	0.015	0.0187	0.0983
66	WARTORN	War Participation 1960-90	0.015	0.3977	0.4922
67	DENS65I	Interior Density	0.015	43.37	88.06

Variables ranked by Posterior Inclusion Probability (PIP) in the benchmark case, defined in equation

(10). Prior inclusion probability for benchmark case equals $\bar{k}/67 = 7/67 = 0.10$.

Table 1: Posterior Inclusion Probabilities (Fixed Prior Model Size)

Rank	Variable	(1)	(2)	(3)		(4)		(5)	(6)
		BACE	Robust	Fixed Degrees of Freedom $\bar{v} = 100$		$\bar{v} = 25$		$\bar{v} = 5$	Bayes
1	East Asian Dummy	0.823	0.876	0.949		0.792		0.816	0.953
2	Primary Schooling Enrollment	0.796	0.828	0.765		0.875		0.796	0.670
3	Investment Price	0.774	0.766	0.714		0.786		0.776	0.485
4	Initial Income (Log GDP in 1960)	0.685	0.672	0.596		0.735		0.727	0.442
5	Fraction of Tropical Area	0.563	0.634	0.689		0.587		0.565	0.501
6	Population Coastal Density	0.428	0.462	0.472		0.579		0.473	0.272
7	Malaria Prevalence	0.252	0.245	0.247		0.196		0.252	0.448
8	Life Expectancy	0.209	0.172	0.152		0.204		0.207	0.179
9	Fraction Confucian	0.206	0.172	0.115		0.126		0.164	0.095
10	Sub-Saharan Africa Dummy	0.154	0.097	0.100		0.217		0.160	0.055
11	Latin American Dummy	0.149	0.097	0.054		0.221		0.173	0.072
12	Fraction GDP in Mining	0.124	0.077	0.062		0.094		0.041	0.041
13	Spanish Colony Dummy	0.123	0.141	0.142		0.090		0.133	0.152
14	Years Open 1950-94	0.119	0.103	0.096		0.070		0.100	0.070
15	Fraction Muslim	0.114	0.106	0.083		0.090		0.098	0.063
16	Fraction Buddhist	0.108	0.082	0.059		0.069		0.098	0.045
17	Ethnolinguistic Fractionalization	0.105	0.108	0.120		0.086		0.107	0.060
18	Government Consumption Share	0.104	0.079	0.133		0.035		0.133	0.077
19	Population Density	0.086	0.067	0.069		0.037		0.037	0.028
20	Real Exchange Rate Distortions	0.082	0.086	0.064		0.074		0.060	0.046
21	Fraction Speaking Foreign Language	0.080	0.081	0.076		0.167		0.073	0.043

Prior inclusion probability with *fixed* prior model size $\bar{k} = 7$ equals $\bar{k}/67 = 7/67 = 0.104$. Benchmark BACE results in column (1) are calculated using 100 million Monte Carlo loops using the stratified sampler. Robust estimates in columns (2)-(8) are calculated using 10 million Monte Carlo (outer) loops, and 200 (inner) Gibbs sampler loops with a burnin of 20 loops. See Computational Appendix A for details.

Table 2: Posterior Mean and Standard Deviations (Fixed Prior Model Size)

Rank	Variable	(1)	(2)	Fixed Degrees of Freedom			(6)	(7)
		BACE	Robust	$\bar{v} = 100$	$\bar{v} = 25$	$\bar{v} = 5$	Bayes	White
1	East Asian Dummy	0.0218	0.0230	0.0224	0.0220	0.0241	0.0249	0.0218
2	Primary Schooling Enrollment	0.0061	0.0063	0.0061	0.0059	0.0062	0.0062	0.0044
3	Investment Price	0.0268	0.0270	0.0264	0.0277	0.0274	0.0248	0.0269
4	Initial Income (Log GDP in 1960)	0.0080	0.0080	0.0080	0.0083	0.0082	0.0082	0.0054
5	Fraction of Tropical Area	-8.3E-05	-8.2E-05	-8.3E-05	-8.6E-05	-8.0E-05	-8.2E-05	-8.4E-05
6	Population Coastal Density	2.5E-05	2.7E-05	2.5E-05	2.5E-05	2.8E-05	2.6E-05	1.0E-05
7	Malaria Prevalence	-0.0085	-0.0082	-0.0083	-0.0086	-0.0082	-0.0081	-0.0085
8	Life Expectancy	0.0029	0.0029	0.0029	0.0028	0.0029	0.0029	0.0018
9	Fraction Confucian	-0.0148	-0.0147	-0.0147	-0.0147	-0.0146	-0.0144	-0.0148
10	Sub-Saharan Africa Dummy	0.0042	0.0042	0.0041	0.0040	0.0041	0.0043	0.0028
11	Latin American Dummy	8.7E-06	8.4E-06	8.6E-06	8.5E-06	8.2E-06	8.6E-06	8.7E-06
12	Fraction GDP in Mining	2.9E-06	2.9E-06	2.9E-06	2.7E-06	2.9E-06	3.0E-06	1.5E-06
13	Spanish Colony Dummy	-0.0158	-0.0164	-0.0163	-0.0154	-0.0163	-0.0176	-0.0157
14	Years Open 1950-94	0.0062	0.0062	0.0065	0.0057	0.0061	0.0056	0.0048
15	Fraction Muslim	0.0008	0.0008	0.0008	0.0007	0.0008	0.0008	0.0008
16	Fraction Buddhist	0.0004	0.0004	0.0004	0.0004	0.0003	0.0004	0.0003
17	Ethnolinguistic Fractionalization	0.0538	0.0543	0.0429	0.0539	0.0538	0.0472	0.0550
18	Government Consumption Share	0.0220	0.0263	0.0227	0.0227	0.0255	0.0239	0.0156
19	Population Density	-0.0153	-0.0141	-0.0105	-0.0163	-0.0168	-0.0122	-0.0150
20	Real Exchange Rate Distortions	0.0071	0.0067	0.0057	0.0060	0.0063	0.0073	0.0057
21	Fraction Speaking Foreign Language	-0.0132	-0.0124	-0.0095	-0.0154	-0.0132	-0.0097	-0.0129
		0.0058	0.0064	0.0056	0.0056	0.0053	0.0056	0.0048
		0.0375	0.0302	0.0327	0.0348	0.0216	0.0343	0.0388
		0.0184	0.0221	0.0196	0.0205	0.0243	0.0194	0.0113
		-0.0109	-0.0117	-0.0102	-0.0094	-0.0099	-0.0111	-0.0107
		0.0050	0.0049	0.0052	0.0058	0.0047	0.0045	0.0035
		0.0124	0.0110	0.0121	0.0116	0.0108	0.0125	0.0121
		0.0063	0.0064	0.0062	0.0065	0.0065	0.0063	0.0039
		0.0125	0.0131	0.0122	0.0116	0.0122	0.0116	0.0127
		0.0061	0.0065	0.0060	0.0070	0.0063	0.0066	0.0045
		0.0109	0.0118	0.0112	0.0109	0.0113	0.0114	0.0073
		-0.0114	-0.0106	-0.0112	-0.0115	-0.0094	-0.0122	-0.0113
		0.0059	0.0059	0.0058	0.0059	0.0059	0.0062	0.0030
		-0.0452	-0.0424	-0.0434	-0.0401	-0.0394	-0.0472	-0.0441
		0.0239	0.0278	0.0235	0.0271	0.0238	0.0232	0.0144
		1.2E-05	1.1E-05	1.1E-05	1.1E-05	9.4E-06	1.1E-05	1.3E-05
		6.7E-06	7.4E-06	6.8E-06	7.2E-06	7.6E-06	7.6E-06	3.4E-06
		-8.0E-05	-7.6E-05	-7.0E-05	-7.2E-05	-7.0E-05	-7.9E-05	-7.8E-05
		4.3E-05	4.1E-05	4.3E-05	4.1E-05	4.2E-05	4.4E-05	2.6E-05
		0.0069	0.0064	0.0067	0.0056	0.0065	0.0066	0.0071
		0.0040	0.0037	0.0040	0.0036	0.0037	0.0040	0.0026

Notes to Table 2: Posterior Mean (13) and Standard Deviation calculated as square root of the variance (14) are *conditional on inclusion* of explanatory variable \mathbf{x}_j . Benchmark BACE results in column (1) and White robust standard errors in column (9) are based on 100 million Monte Carlo loops using the stratified sampler. Robust estimates in columns (2)-(8) are based on 10 million Monte Carlo (outer) loops, and 200 (inner) Gibbs sampler loops with a burnin of 20 loops. See Computational Appendix A for details.

Table 3: Posterior Inclusion Probabilities (Random Prior Model Size)

Rank	Variable	(1)	(2)	(3)		(4)	(5)	(6)
		BACE	Robust	Fixed Degrees of Freedom		$\bar{v} = 25$	$\bar{v} = 5$	Bayes
1	East Asian Dummy	0.778	0.892	0.800	0.766	0.782	0.971	0.971
2	Primary Schooling Enrollment	0.793	0.778	0.776	0.775	0.847	0.607	0.607
3	Investment Price	0.786	0.775	0.755	0.774	0.813	0.367	0.367
4	Initial Income (Log GDP in 1960)	0.722	0.700	0.715	0.696	0.673	0.325	0.325
5	Fraction of Tropical Area	0.530	0.625	0.534	0.533	0.589	0.427	0.427
6	Coastal Population Density	0.416	0.472	0.439	0.390	0.469	0.225	0.225
7	Malaria Prevalence	0.239	0.267	0.253	0.224	0.209	0.536	0.536
8	Life Expectancy	0.242	0.222	0.267	0.271	0.180	0.134	0.134
9	Fraction Confucian	0.270	0.166	0.246	0.271	0.244	0.069	0.069
10	Sub-Saharan Africa Dummy	0.208	0.098	0.194	0.205	0.159	0.044	0.044
11	Latin American Dummy	0.189	0.087	0.174	0.147	0.166	0.054	0.054
12	Fraction GDP in Mining	0.203	0.139	0.216	0.248	0.092	0.023	0.023
13	Spanish Colony Dummy	0.127	0.110	0.115	0.144	0.122	0.167	0.167
14	Years Open 1950-94	0.123	0.132	0.130	0.129	0.132	0.059	0.059
15	Fraction Muslim	0.150	0.144	0.145	0.166	0.136	0.042	0.042
16	Fraction Buddhist	0.142	0.101	0.120	0.127	0.158	0.034	0.034
17	Ethnolinguistic Fractionalization	0.115	0.130	0.106	0.127	0.129	0.045	0.045
18	Government Consumption Share	0.127	0.146	0.121	0.095	0.107	0.078	0.078
19	Population Density	0.128	0.103	0.108	0.110	0.104	0.018	0.018
20	Real Exchange Rate Distortions	0.116	0.089	0.118	0.157	0.074	0.046	0.046
21	Fraction Speaking Foreign Language	0.111	0.082	0.093	0.140	0.098	0.036	0.036

Prior inclusion probability with *fixed* prior model size $\bar{k} = 7$ equals $\bar{k}/67 = 7/67 = 0.104$. Benchmark BACE results in column (1) are calculated using 100 million Monte Carlo loops using the stratified sampler. Robust estimates in columns (2)-(8) are calculated using 10 million Monte Carlo (outer) loops, and 200 (inner) Gibbs sampler loops with a burnin of 20 loops. See Computational Appendix A for details.

Table 4: Posterior Mean and Standard Deviations (Random Prior Model Size)

Rank	Variable	(1)		(2)		Fixed Degrees of Freedom			(5)	(6)	(7)
		BACE	Robust	$\bar{v} = 100$			$\bar{v} = 25$			Bayes	White
1	East Asian Dummy	0.0213	0.0228	0.0217	0.0223	0.0238	0.0223	0.0238	0.0260	0.0213	0.0213
2	Primary Schooling Enrollment	0.0066	0.0067	0.0067	0.0063	0.0065	0.0063	0.0065	0.0063	0.0049	0.0049
3	Investment Price	0.0269	0.0267	0.0267	0.0270	0.0265	0.0270	0.0265	0.0239	0.0267	0.0267
4	Initial Income (Log GDP in 1960)	0.0080	0.0082	0.0083	0.0078	0.0082	0.0078	0.0082	0.0080	0.0055	0.0055
5	Fraction of Tropical Area	-8.4E-05	-8.1E-05	-8.3E-05	-8.4E-05	-8.1E-05	-8.4E-05	-8.1E-05	-8.1E-05	-8.1E-05	-8.6E-05
6	Coastal Population Density	2.5E-05	2.7E-05	2.5E-05	2.6E-05	2.9E-05	2.6E-05	2.9E-05	2.6E-05	2.6E-05	1.2E-05
7	Malaria Prevalence	-0.0088	-0.0086	-0.0089	-0.0091	-0.0084	-0.0091	-0.0084	-0.0079	-0.0090	-0.0090
8	Life Expectancy	0.0030	0.0031	0.0030	0.0031	0.0029	0.0030	0.0029	0.0029	0.0018	0.0018
9	Fraction Confucian	0.0143	0.0144	0.0144	0.0140	0.0143	0.0140	0.0143	0.0145	0.0144	0.0144
10	Sub-Saharan Africa Dummy	0.0045	0.0044	0.0045	0.0044	0.0044	0.0044	0.0044	0.0043	0.0031	0.0031
11	Latin American Dummy	8.5E-06	8.4E-06	8.3E-06	8.4E-06	8.0E-06	8.4E-06	8.0E-06	8.4E-06	8.5E-06	8.5E-06
12	Fraction GDP in Mining	3.0E-06	3.0E-06	3.0E-06	2.9E-06	3.0E-06	2.9E-06	3.0E-06	2.9E-06	1.6E-06	1.6E-06
13	Spanish Colony Dummy	-0.0155	-0.0153	-0.0161	-0.0161	-0.0158	-0.0161	-0.0158	-0.0190	-0.0151	-0.0151
14	Years Open 1950-94	0.0067	0.0063	0.0066	0.0062	0.0064	0.0062	0.0064	0.0056	0.0051	0.0051
15	Fraction Muslim	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008
16	Fraction Buddhist	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0002	0.0002
17	Ethnolinguistic Fractionalization	0.0532	0.0480	0.0527	0.0563	0.0543	0.0527	0.0563	0.0515	0.0543	0.0543
18	Government Consumption Share	0.0217	0.0244	0.0219	0.0221	0.0256	0.0221	0.0256	0.0256	0.0122	0.0122
19	Population Density	-0.0151	-0.0124	-0.0153	-0.0143	-0.0161	-0.0143	-0.0161	-0.0161	-0.0146	-0.0146
20	Real Exchange Rate Distortions	0.0068	0.0065	0.0071	0.0065	0.0066	0.0065	0.0066	0.0068	0.0043	0.0043
21	Fraction Speaking Foreign Language	-0.0129	-0.0106	-0.0128	-0.0137	-0.0126	-0.0137	-0.0126	-0.0088	-0.0136	-0.0136
		0.0059	0.0062	0.0059	0.0061	0.0060	0.0061	0.0060	0.0050	0.0037	0.0037
		0.0439	0.0325	0.0434	0.0450	0.0322	0.0450	0.0322	0.0323	0.0534	0.0534
		0.0200	0.0211	0.0201	0.0203	0.0253	0.0203	0.0253	0.0195	0.0155	0.0155
		-0.0101	-0.0102	-0.0099	-0.0105	-0.0107	-0.0105	-0.0107	-0.0112	-0.0100	-0.0100
		0.0052	0.0054	0.0053	0.0050	0.0050	0.0050	0.0050	0.0042	0.0039	0.0039
		0.0116	0.0115	0.0114	0.0101	0.0100	0.0101	0.0100	0.0129	0.0116	0.0116
		0.0063	0.0067	0.0062	0.0059	0.0060	0.0059	0.0060	0.0064	0.0041	0.0041
		0.0126	0.0123	0.0126	0.0133	0.0129	0.0133	0.0129	0.0114	0.0127	0.0127
		0.0063	0.0063	0.0063	0.0067	0.0062	0.0067	0.0062	0.0064	0.0046	0.0046
		0.0211	0.0196	0.0199	0.0228	0.0219	0.0228	0.0219	0.0215	0.0200	0.0200
		0.0104	0.0115	0.0106	0.0106	0.0115	0.0106	0.0115	0.0130	0.0050	0.0050
		-0.0109	-0.0104	-0.0105	-0.0099	-0.0095	-0.0099	-0.0095	-0.0122	-0.0109	-0.0109
		0.0059	0.0058	0.0059	0.0059	0.0059	0.0059	0.0059	0.0064	0.0031	0.0031
		-0.0436	-0.0421	-0.0417	-0.0418	-0.0385	-0.0418	-0.0385	-0.0488	-0.0432	-0.0432
		0.0256	0.0241	0.0273	0.0246	0.0254	0.0246	0.0254	0.0244	0.0161	0.0161
		1.3E-05	1.2E-05	1.3E-05	1.3E-05	1.0E-05	1.3E-05	1.0E-05	1.1E-05	1.3E-05	1.3E-05
		6.9E-06	7.4E-06	7.2E-06	7.1E-06	7.7E-06	7.1E-06	7.7E-06	8.2E-06	3.8E-06	3.8E-06
		-8.1E-05	-7.3E-05	-8.1E-05	-8.2E-05	-7.0E-05	-8.2E-05	-7.0E-05	-7.6E-05	-7.9E-05	-7.9E-05
		4.4E-05	4.3E-05	4.4E-05	3.9E-05	4.1E-05	3.9E-05	4.1E-05	4.3E-05	2.7E-05	2.7E-05
		0.0074	0.0068	0.0073	0.0072	0.0066	0.0072	0.0066	0.0065	0.0086	0.0086
		0.0046	0.0040	0.0041	0.0035	0.0039	0.0035	0.0039	0.0046	0.0021	0.0021

Notes to Table 4: Posterior Mean (13) and Standard Deviation calculated as square root of the variance (14) are *conditional on inclusion* of explanatory variable \mathbf{x}_j . Benchmark BACE results in column (1) and White robust standard errors in column (9) are based on 100 million Monte Carlo loops using the stratified sampler. Robust estimates in columns (2)-(8) are based on 10 million Monte Carlo (outer) loops, and 200 (inner) Gibbs sampler loops with a burnin of 20 loops. See Computational Appendix A for details.

Figure 2: BACE (BLUE) vs. Robust (RED) Distributions (Random d.f. $v_0 = 25$).

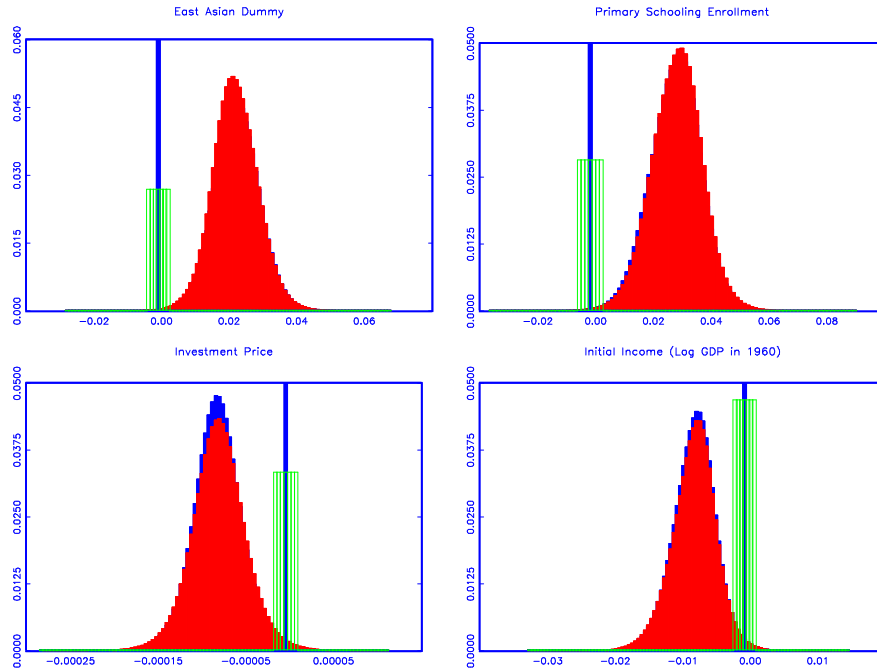


Figure 3: BACE (BLUE) vs. Robust (RED) Distributions (Random d.f. $v_0 = 25$).

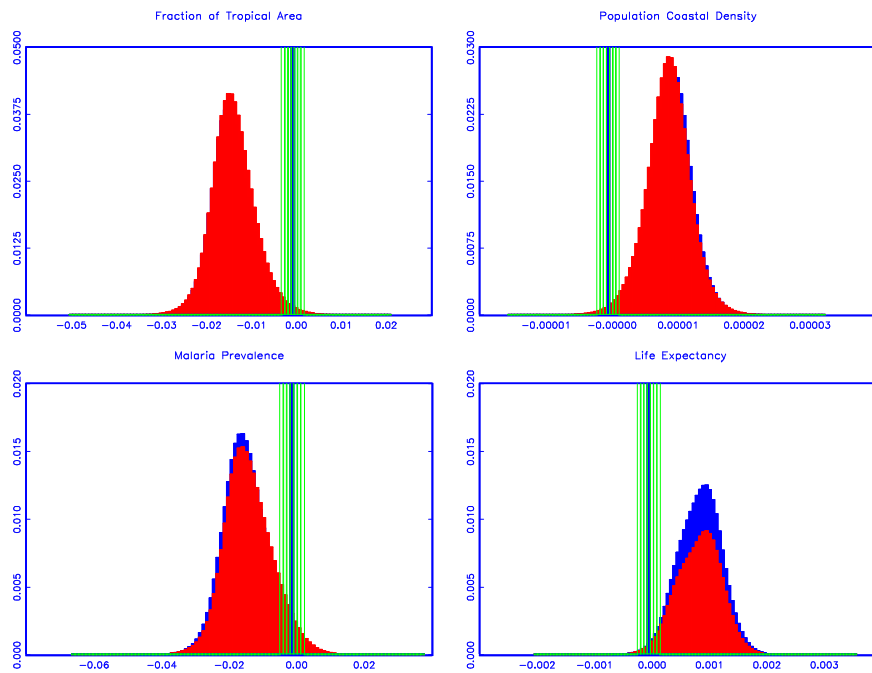


Figure 4: BACE (BLUE) vs. Robust (RED) Distributions (Random d.f. $v_0 = 25$).

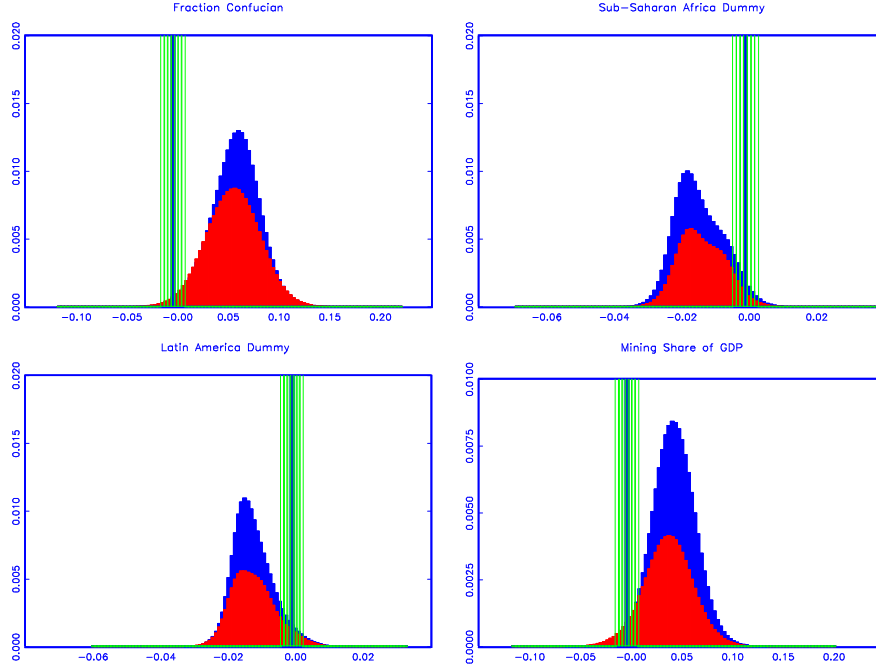


Figure 5: BACE (BLUE) vs. Robust (RED) Distributions (Random d.f. $v_0 = 25$).

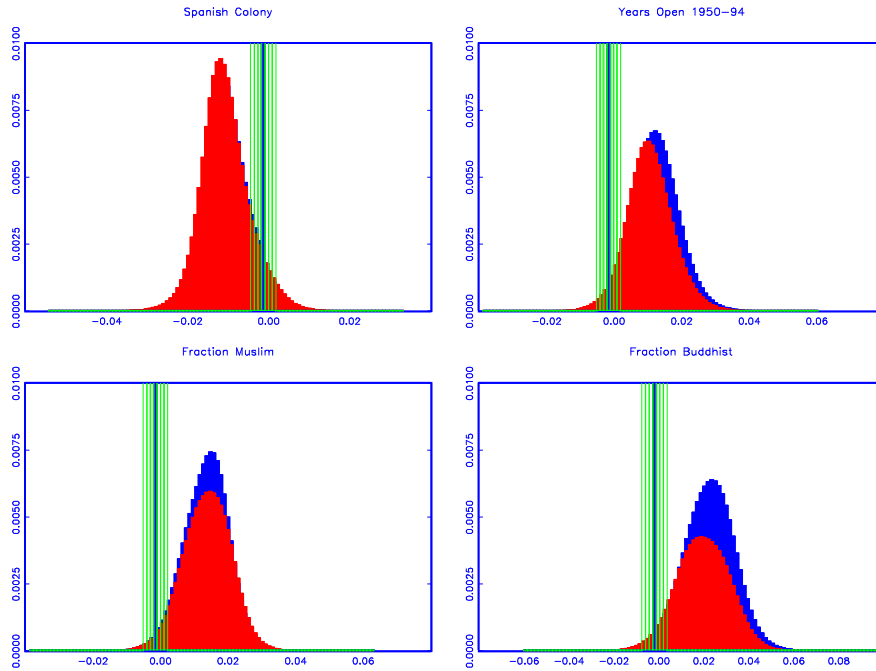


Figure 6: BACE (BLUE) vs. Robust (RED) Distributions (Random d.f. $v_0 = 25$).

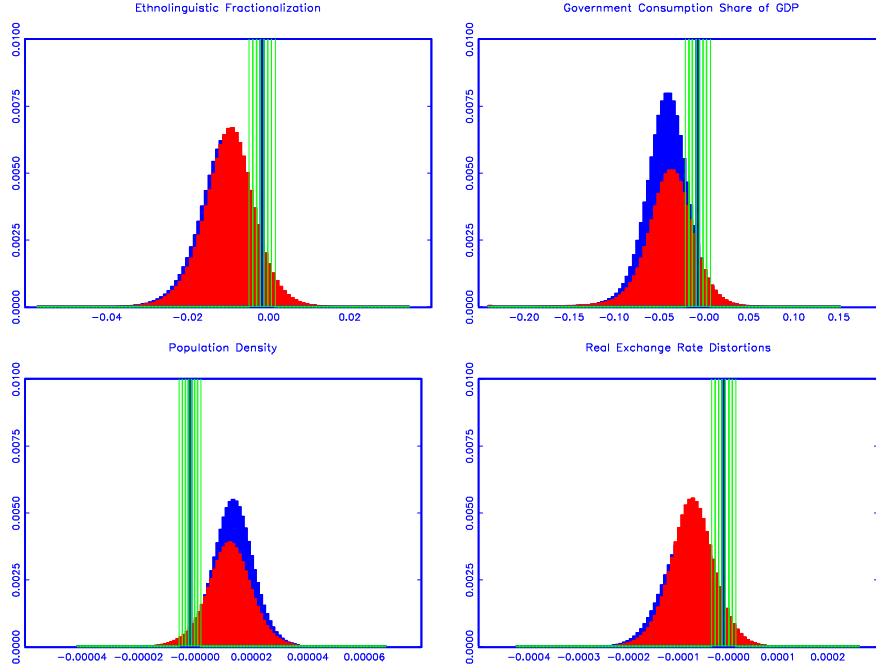


Figure 7: BACE (BLUE) vs. Robust (RED) Distributions (Random d.f. $v_0 = 25$).

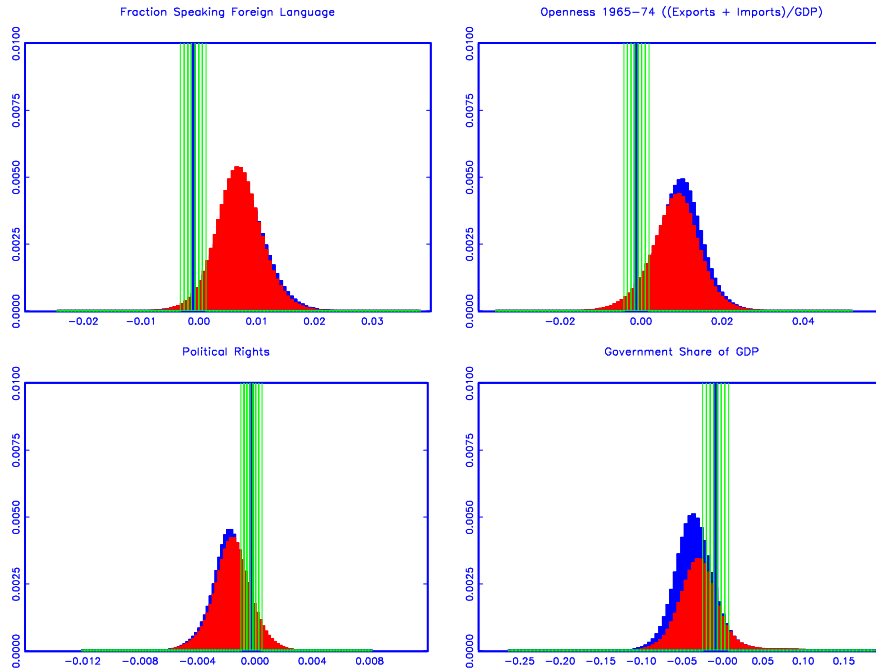


Figure 8: Posterior Error Variances $\Omega = \text{diag}(\omega_i)$.

