# `bdsm`: Bayesian dynamic systems modelling. Bayesian model averaging for dynamic panels with weakly exogenous regressors


Mateusz Wyszyński
Lazarski University
University of Warsaw

Krzysztof Beck
Lazarski University

Marcin Dubel
Lazarski University


April 11, 2025


**Abstract**

This manuscript introduces the `bdsm` package, which enables Bayesian model averaging for dynamic panels with weakly exogenous regressors—a methodology developed by Moral-Benito (2016). The package allows researchers to simultaneously address model uncertainty and reverse causality. The manuscript includes a hands-on tutorial accessible to users unfamiliar with this approach. In addition to calculating the model space (using parallel computing) and providing key BMA statistics, the package offers flexible options for specifying model priors, including a dilution prior that accounts for multicollinearity. It also provides graphical tools for visualizing prior and posterior model probabilities, as well as functions for plotting histograms and kernel densities of the estimated coefficients. Furthermore, the package enables researchers to compute jointness measures and perform Bayesian model selection to examine the most probable models based on posterior model probabilities.


## Contents

# 1 Introduction

Since the seminal works of Leamer (1978); Leamer & Leonard (1981); Leamer (1983, 1985), there has been an increased focus on reporting the fragility of regression estimates. Leamer (1983) proposed Extreme Bounds Analysis (EBA) as a remedy for addressing the sensitivity of empirical research findings[1]. In economics, growth regressions (Barro, 1991) became a central focus of research on economic growth during the 1990s. However, the credibility of these results was challenged when Levine & Renelt (1992) applied EBA to cross-country economic growth data. The authors found that investment as a share of GDP was the only variable robust to changes in model specification. In response, EBA was criticized for being too stringent[2], leading to the proposal of alternative approaches (Sala-I-Martin, 1997).

Bayesian model averaging (BMA) emerged as a preferred method during a period when studies of economic growth advanced alongside methodological innovations (Fernández et al., 2001a,b; Sala-I-Martin et al., 2004; Eicher et al., 2007; Ley & Steel, 2012; Moser & Hofmarcher, 2014; Fernández et al., 2001a,b; Arin et al., 2019). As a result, BMA became a widely used technique for assessing the robustness of regressors in economics[3] (e.g., Liu & Maheu (2009); Ductor & Leiva-Leon (2016); Figini & Giudici (2017); Beck (2022); D'Andrea (2022); Horvath et al. (2024)), as well as in other fields (e.g., Sloughter et al. (2013); Baran & Möller (2015); Aller et al. (2021); Guliyev (2024); Payne et al. (2024)). Moreover, the growing interest in BMA was fueled by the availability of R packages such as `BMA` (Raftery et al., 2005), `BAS` (Clyde et al., 2011), and `BMS` (Feldkircher & Zeugner, 2015), along with the `gretl` BMA package developed by Błażejowski & Kwiatkowski (2015).

The primary issue with the Bayesian model averaging in the aforementioned studies was its reliance on the assumption of exogenous regressors. In many contexts, particularly in economics, this premise is unsuitable. Instead, the assumption of endogenous variables within a simultaneous equations framework is more fitting. Consequently, a new line of research relaxed the assumption of exogenous regressors Lenkoski et al. (2014); León-González & Montolio (2015); Mirestean & Tsangarides (2016); Moral-Benito (2016); Chen et al. (2018). However, these methods have not found their way into mainstream research. The code to implement them is only available upon request from the authors and is provided exclusively for `MATLAB` and `GAUSS`.

The `bdsm` package was developed to address this gap. It offers tools for performing Bayesian model averaging on dynamic panels with weakly exogenous regressors. As a result, it enables researchers to address both model uncertainty and reverse causality. The core of the code is based on the methodological approach developed by Moral-Benito (2012, 2013, 2016). While the main aspects of the method are described in the manuscript, interested readers should refer to the original articles for further details. In addition to the key features developed by Moral-Benito (2016), the `bdsm` package offers a wide range of additional functionalities. The package enables users to employ flexible model prior options, along with a dilution prior, which helps account for multicollinearity. The `bdsm` package provides users with graphical options for plotting prior and posterior model probabilities across model sizes and the model space. Additionally, users can utilize Bayesian model selection to thoroughly examine the best models based on posterior model probability. The package calculates jointness measures developed by Doppelhofer & Weeks (2009); Ley & Steel (2007); Hofmarcher et al. (2018). Finally, it offers users the option to plot histograms or kernel densities of the estimated coefficients for the examined regressors.

The remainder of the manuscript is structured as follows. Section 2 describes the dynamic panel setup considered by Moral-Benito (2013) and outlines the Bayesian model averaging approach used in the package. Data preparation is detailed in Section 3, while Section 4 addresses the estimation of the model space. Section 5 provides an overview of the `bdsm` functions related to performing Bayesian model averaging, calculating jointness measures, and presenting the estimation results. The details of the model prior choices are described in Section 6. Finally, Section 7 offers some concluding remarks.

[1] Hlavac (2016) developed an R package for EBA.
[2] Granger & Uhlig (1990) proposed a less restrictive variant of EBA.
[3] For a detailed review of BMA applications in economics, see Moral-Benito (2015); Steel (2020).

# 2 Model setup and Bayesian model averaging

This section outlines the model setup, describes the approach to Bayesian model averaging implemented in the package, summarizes the main BMA statistics, and discusses model priors and jointness measures.

## 2.1 Model setup

Moral-Benito (2016) considers the following model specification:

$$y_{it} = \alpha y_{it-1} + \beta x_{it} + \eta_i + \zeta_t + \upsilon_{it} \tag{1}$$

where $y_{it}$ is the dependent variable, $i$ $(= 1, ..., N)$ indexes entity (ex. country), $t$ $(= 1, ..., T)$ indexes time, $x_{it}$ is a matrix of growth determinants, $\beta$ is a parameter vector, $\eta_i$ is an entity specific fixed effect, $\zeta_t$ is a period-specific shock and $\upsilon_{it}$ is a shock to the dependent variable. To address the issue of reverse causality the model is build on the assumption of weak exogeneity, that can be formalized as

$$\mathbb{E}(\upsilon_{i,t}|y_t^{t-1}, x_i^t, \eta_i) = 0 \tag{2}$$

where $y_t^{t-1} = (y_{i,0}, ..., y_{i,t-1})'$ and $x_i^t = (x_{i,0}, ..., x_{i,t})'$. Accordingly, weak exogeneity implies that the current values of the regressors, lagged dependent variable, and fixed effects are uncorrelated with the current shocks, while they are all allowed to be correlated with each other at the same time. On the assumption of weakly exogenous regressors, Moral-Benito (2013) augmented equation (1) with additional reduced-form equations capturing the unrestricted feedback process:

$$x_{it} = \gamma_{t0} y_{i0} + ... + \gamma_{tt-1} y_{it-1} + \Lambda_{t1} x_{i1} + ... + \Lambda_{tt-1} x_{it-1} + c_t \eta_i + \vartheta_{it} \tag{3}$$

where $t = 2, \ldots, T$; $c_t$ is the $k \times 1$ vector of parameters. For $h < t$, $\gamma_{th}$ is a $k \times 1$ vector $(y_{th}^1, \ldots, y_{th}^k)'$ $h = 0, \ldots, T-1$; $\Lambda_{th}$ is a $k \times k$ matrix of parameters, and $\vartheta_{it}$ is a $k \times 1$ of prediction errors. The initial observations are defined with

$$y_{i0} = c_0 \eta_i + \upsilon_{it} \tag{4}$$

$$x_{i1} = \gamma_{10} y_{i0} + c_1 \eta_i + \vartheta_{it} \tag{5}$$

where $c_0$ is a scalar, $c_1$ and $\gamma_{10}$ are $k \times 1$ vectors and $\eta_i$ are the individual effects. The mean vector and the covariance matrix of the joint distribution of the initial observations and the individual effects are unrestricted.[4] For the model setup given in equations (1) and (3-5), Moral-Benito (2013) derived the log-likelihood function:

$$\log f(data|\theta) \propto \frac{N}{2} \log \det(B^{-1} D \Sigma D' B'^{-1}) - \frac{1}{2} \sum_{i=1}^{N} \{R_i' (B^{-1} D \Sigma D' B'^{-1})^{-1} R_i\} \tag{6}$$

where $\theta$ denotes parameters to be estimated, $R_i = (y_{io}, x_{i1}', y_{i1}, \ldots, x_{iT}', y_{iT})'$ are vectors of observed variables, and $\Sigma = diag[\sigma_\eta^2, \sigma_{\upsilon_0}^2, \Sigma_{\vartheta_1}, \sigma_{\upsilon_1}^2, ..., \Sigma_{\vartheta_T}, \sigma_{\upsilon_T}^2]$ is the block-diagonal variance-covariance matrix. Matrix B is given by:

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\gamma_{10} & I_k & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\alpha & -\beta' & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\gamma_{20} & -\Lambda_{21} & -\gamma_{21} & I_k & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\alpha & -\beta' & 1 & \cdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 & 0 \\ -\gamma_{T0} & -\Lambda_{T1} & -\gamma_{T1} & -\Lambda_{T2} & -\gamma_{T2} & \cdots & -\gamma_{TT-1} & I_k & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & -\alpha & -\beta' & 1 \end{bmatrix} \tag{7}$$

and matrix D is given by:

$$D = \begin{bmatrix} (c_0 & c_1' & 1 & c_2' & 1 & \ldots & c_T' & 1)' & I_{T(k+1)+1} \end{bmatrix}. \tag{8}$$

The model setup in equations (1) and (3-5) requires that in addition to the parameters of interest $\alpha$ and $\beta$, the parameters $\gamma_{ij}$ and $\Lambda_{km}$ need to be estimated. To make the optimization

[4] The method outperforms the Arellano–Bond estimator (Moral-Benito et al., 2019).

of likelihood computationally feasible, Moral-Benito (2013) developed Simultaneous Equations Model (SEM) setup where the parameters of non-central interest are incorporated in the variance-covariance matrix. In the SEM setup, the model is defined by $1 + (T-1)k + T$ equations:

$$\begin{cases} \eta_i = \phi_i y_{i0} + x'_{i1}\phi_1 + \epsilon_i & \\ x_{it} = \pi_{t0} y_{i0} + \pi_{t1} x_{i1} + \pi^w_t x_{i1} + \xi_{it}, & t = 2, ..., T \\ y_{it} = \alpha y_{it-1} + x'_{it}\beta + \phi_0 y_{i0} + x'_{i1}\phi_1 + w'_i\delta + \epsilon_i + v_{it}, & t = 1, ..., T \end{cases}$$

This setup can be rewritten in a matrix form:

$$BR_i = Cz_i + U_i, \tag{9}$$

where:

$$z_i = [y_{i0}, x'_{i1}, w'_i]' \tag{10}$$

is the vector of strictly exogenous variables,

$$R_i = [y_{i1}, y_{i2}, ..., y_{iT}, x'_{i2}, x'_{i3}, ..., x'_{iT}]', \tag{11}$$

$$U_i = [\epsilon_i + v_{i1}, \epsilon_i + v_{i2}, ..., \epsilon_i + v_{iT}, \xi'_{i2}, \xi'_{i3}, ..., \xi'_{iT}]' \tag{12}$$

and matrices $B$ and $C$ contain coefficients $\alpha$, $\beta$, $\phi_0$, $\phi_1$. Since these matrices are not connected to the error, we simply note that they are defined in such a way that the equation (9) is equivalent to the SEM setup. The main difference of the SEM setup is that equations for $x_{it}$ now depend only on $y_{i0}$ and $x_{i1}$ and not on $y_{is}$ and $x_{is}$ for other periods $s$.

Following Moral-Benito (2013), we can then define the likelihood function as:

$$L \propto -\frac{N}{2} \log\det \Omega(\theta) - \frac{1}{2} tr\{\Omega(\theta)^{-1}(R - Z\Pi(\theta))'(R - Z\Pi(\theta))\} \tag{13}$$

where $R$ and $Z$ are matrices containing vectors $R_i$ and $z_i$ respectively and:

$$\Pi(\theta) = B^{-1}C \tag{14}$$

$$U^*_i(\theta) = B^{-1}U_i \tag{15}$$

$$\Omega(\theta) = Var(U^*_i) = B^{-1} \cdot Var(U_i) \cdot B'^{-1} = B^{-1}\Sigma B'^{-1} \tag{16}$$

It is possible to find analytical solution for MLE for some of the parameters. Then the formula for the likelihood function can be simplified to:

$$L(\theta) \propto -\frac{N}{2} \log\det \Sigma_{11} - \frac{1}{2} tr\{\Sigma^{-1}_{11} U'_1 U_1\} - \frac{N}{2} \log\det(\frac{H}{N}) \tag{17}$$

where $U_1$ is a matrix of errors connected only to dependent variables, $\Sigma_{11}$ is a part of the $\Sigma$ matrix:

$$\Sigma = var(U_i) = var\begin{pmatrix} U_{i1} \\ U_{i2} \end{pmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}. \tag{18}$$

and $H = (R_2 + U_1F_{12})'Q(R_2 + U_1F_{12})$ with $R_2$ being a matrix of regressor vectors $[x'_{i2}, x'_{i3}, ..., x'_{iT}]$ and $F_{12} = -\Sigma^{-1}_{11}\Sigma_{12}$.

## 2.2 Bayesian model averaging

Given the likelihood function in (17), henceforth denoted as $L(\text{data}|\theta_i, M_i)$ for a specific model $i$, it is possible to utilize Bayesian model averaging[5] (BMA). To achieve that, we first estimate all possible variants of equation:

$$Y = f(X, \theta, v) \tag{19}$$

where $Y$ is a vector of dependent variable, $X$ is a matrix of potential determinants, $\theta$ is a parameter vector, and $v$ is a stochastic term. All the variants include a lagged dependent variable; therefore, with $K$ regressors, there are $2^K$ possible models that

[5] For an introduction to BMA see Raftery (1995); Raftery et al. (1997); Kass & Raftery (1995); Doppelhofer & Weeks (2009); Amini & Parmeter (2011); Beck (2017); Fragoso et al. (2018).

can be estimated. Each of these models can be assigned a posterior model probability; however, the marginal (integrated) likelihood, $L(\text{data}|M_i)$, must first be computed. Moral-Benito (2012) utilizes approach of developed by Raftery (1995) and Sala-I-Martin et al. (2004) based on the Bayesian information criterion (BIC) approximation.

The Bayes factor for models $M_i$ and $M_i$, $B_{ij} = \frac{L(\text{data}|M_i)}{L(\text{data}|M_j)}$, can be approximated using Schwartz criterion:

$$S = \log L(\text{data} \mid \hat{\theta}_i, M_i) - \log L(\text{data}|\hat{\theta}_j, M_j) - \frac{k_i - k_j}{2} \log(N) \tag{20}$$

where $L(\text{data}|\hat{\theta}_i, M_i)$ and $L(\text{data}|\hat{\theta}_j, M_j)$ are the maximum likelihood values for models $i$ and $j$, respectively. The terms $k_i$ and $k_j$ denote the number of regressors in models $i$ and $j$. Bayesian information criterion is given by:

$$BIC = -2S = -2 \log B_{ij}. \tag{21}$$

Given null model $M_0$

$$B_{ij} = \frac{L(y|M_i)}{L(y|M_j)} = \frac{\frac{L(y|M_i)}{L(y|M_0)}}{\frac{L(y|M_j)}{L(y|M_0)}} = \frac{B_{i0}}{B_{j0}} = \frac{B_{0j}}{B_{0i}} \tag{22}$$

and

$$2 \log B_{ij} = 2[\log B_{0j} - \log B_{0i}] = BIC_j - BIC_i. \tag{23}$$

The posterior model probability (PMP) of model $j$ given the data is

$$\mathbb{P}(M_j|y) = \frac{L(\text{data}|M_j)\mathbb{P}(M_j)}{\sum_{i=1}^{2^K} L(\text{data}|M_i)\mathbb{P}(M_i)} \tag{24}$$

where $\mathbb{P}(M_j)$ denotes prior model probability. In other words, the PMP represents the share of model $j$ in the total posterior probability mass. Combining, equations (21-24) we get:

$$\begin{aligned} \mathbb{P}(M_j|y) &= \frac{L(\text{data}|M_j)\mathbb{P}(M_j)}{\sum_{i=1}^{2^K} L(\text{data}|M_i)\mathbb{P}(M_i)} = \frac{\frac{L(\text{data}|M_j)}{L(\text{data}|M_0)} L(\text{data}|M_0)\mathbb{P}(M_j)}{\sum_{i=1}^{2^K} \frac{L(\text{data}|M_i)}{L(\text{data}|M_0)} L(\text{data}|M_0)\mathbb{P}(M_i)} \\ &= \frac{B_{j0} L(\text{data}|M_0)\mathbb{P}(M_j)}{\sum_{i=1}^{2^K} B_{i0} L(\text{data}|M_0)\mathbb{P}(M_i)} = \frac{L(\text{data}|M_0) B_{j0}\mathbb{P}(M_j)}{L(\text{data}|M_0)\sum_{i=1}^{2^K} B_{i0}\mathbb{P}(M_i)} = \frac{B_{j0}\mathbb{P}(M_j)}{\sum_{i=1}^{2^K} B_{i0}\mathbb{P}(M_i)} \end{aligned} \tag{25}$$

Finally, using the result that

$$B_{j0} = \exp\left(-\frac{1}{2} BIC_j\right) \tag{26}$$

we can calculate posterior model probability as

$$\mathbb{P}(M_j|\text{data}) = \frac{\exp\left(-\frac{1}{2} BIC_j\right)\mathbb{P}(M_j)}{\sum_{i=1}^{2^K} \exp\left(-\frac{1}{2} BIC_i\right)\mathbb{P}(M_i)}. \tag{27}$$

## 2.3 BMA statistics

With PMPs, we can calculate useful BMA statistics. Let's denote by $\pi_k$ the random variable which is equal to one if the $k^{th}$ regressor should be considered as the determinant of the dependent variable. The posterior inclusion probability (PIP) for the regressor is given by:

$$\mathbb{P}(\pi_k = 1|\text{data}) = \sum_{j=1}^{2^K} \mathbb{1}(k^{th} \text{ regressor is in model } M_j) \cdot \mathbb{P}(M_j|\text{data}) \tag{28}$$

where the indicator function $\mathbb{1}$ is equal to one if the regressor is part of the model $M_j$ and zero otherwise. In other words, the PIP tells us how likely it is that the given regressor has impact on the variable of interest.

Another interesting statistic is the posterior mean (PM) of a given parameter $\beta$. Let's denote by $\pi_\beta$ the random variable which is equal to one if the given parameter is present in the model, and zero otherwise. The posterior mean of $\beta$ is given by:

$$\mathbb{E}(\beta|\text{data}) = \sum_{j=1}^{2^K} \widehat{\beta}_j \cdot \mathbb{P}(M_j, \pi_\beta = 1|\text{data}) \tag{29}$$

where $\widehat{\beta}_j$ is the value of the coefficient $\beta$ in model $j$. It tells us what is the mean (or expected) value for the parameter taking into account all considered models. Note that if $\beta$ is not present in the given model $j$ we can assign any value to $\widehat{\beta}_j$, because the probability $\mathbb{P}(M_j, \pi_\beta = 1|\text{data})$ wil be zero anyway.

The posterior variance of the parameter $\beta$ is equal to:

$$\begin{aligned} Var(\beta|\text{data}) = &\sum_{j=1}^{2^K} Var(\beta_j|\text{data}, M_j) \cdot \mathbb{P}(M_j, \pi_\beta = 1|\text{data}) \\ &+ \sum_{j=1}^{2^K} \left[\widehat{\beta}_j - \mathbb{E}(\beta|\text{data})\right]^2 \cdot \mathbb{P}(M_j, \pi_\beta = 1|\text{data}) \end{aligned} \tag{30}$$

where $Var(\beta_j|\text{data}, M_j)$ denotes the conditional variance of the coefficient $\beta$ in model $M_j$ (in other words assuming that the model $M_j$ is the true model). Posterior standard deviation (PSD) of $\beta$ is then defined as the square root of the variance:

$$SD(\beta|\text{data}) = \sqrt{Var(\beta|\text{data})} \tag{31}$$

Alternatively, one might be interested in the values of the mean and variance on the condition of inclusion of a given parameter, i.e. assuming that it is definitely a part of the model. Note that this is usually determined by the presence of a related regressor. The conditional posterior mean (PMcon) for a parameter $\beta$ is given by:

$$\mathbb{E}(\beta|\pi_\beta = 1, \text{data}) = \frac{\mathbb{E}(\beta|\text{data})}{\mathbb{P}(\pi_\beta = 1|\text{data})}. \tag{32}$$

Similarly, the conditional variance is:

$$Var(\beta|\pi_\beta = 1, \text{data}) = \frac{Var(\beta|\text{data}) + \mathbb{E}(\beta|\text{data})^2}{\mathbb{P}(\pi_\beta = 1|\text{data})} - \mathbb{E}(\beta|\pi_\beta = 1, \text{data})^2 \tag{33}$$

and so the conditional standard deviation (PSDcon) is:

$$SD(\beta|\pi_k = 1, \text{data}) = \sqrt{Var(\beta|\pi_k = 1, \text{data})} \tag{34}$$

The BMA statistics allow the assessment of the robustness of the examined regressors. Raftery (1995), classifies a variable as weak, positive, strong, and very strong when the posterior inclusion probability (PIP) is between 0.5 and 0.75, between 0.75 and 0.95, between 0.95 and 0.99, and above 0.99, respectively. Raftery (1995) also refers to the variable as robust when the absolute value of the ratio of posterior mean (PM) to posterior standard deviation (PSD) is above 1, indicating that the regressor improves the power of the regression. Masanjala & Papageorgiou (2008) propose a more stringent criterion, where they require the statistic to be higher than 1.3, while Sala-I-Martin et al. (2004) argue for 2, corresponding to 90% and 95%, respectively.

## 2.4 Model priors and jointness

To perform BMA one needs to specify prior model probability[6]. The package offers two main options. The first is binomial model prior (Sala-I-Martin et al., 2004):

$$\mathbb{P}(M_j) = (\frac{EMS}{K})^{k_j}(1 - \frac{EMS}{K})^{K-k_j} \tag{35}$$

where $EMS$ is the expected model size and $k_j$ is a number of regressors in model $j$. If $EMS = \frac{K}{2}$, the binomial model prior simplifies to a uniform model prior with $\mathbb{P}(M_j) = \frac{1}{2^K}$ for every $j$, meaning that all models are assumed to have equal probabilities. The second is binomial-beta model prior Ley & Steel (2009) given by:

$$\mathbb{P}(M_j) \propto \Gamma(1 + k_j) \cdot \Gamma(\frac{K - EMS}{EMS} + K - k_j). \tag{36}$$

where $\Gamma$ is the gamma function. In the context of the binomial-beta prior $EMS = \frac{K}{2}$ corresponds to equal probabilities on model sizes.

In order to account for potential multicolinearity between regressors one can use dilution prior introduced by George (2010). The dilution prior involves augmenting the model prior (binomial or binomial-beta) with a function that accounts for multicollinearity:

$$\mathbb{P}_D(M_j) \propto \mathbb{P}(M_j)|COR_j|^{\omega} \tag{37}$$

where $\mathbb{P}_D(M_j)$ is the diluted model prior, $|COR_j|$ is the determinant of the correlation matrix of regressors in model $j$, and $\omega$ is the dilution parameter. The lower the correlation between regressors, the closer $|COR_j|$ is to one, resulting in a smaller degree of dilution.

To determine whether regressors are substitutes or complements, various authors have developed jointness measures[7]. Assuming two different covariates $a$ and $b$, let $\mathbb{P}(a \cap b)$ be the posterior probability of the inclusion of both variables, $\mathbb{P}(\overline{a} \cap \overline{b})$ the posterior probability of the exclusion of both variables, $\mathbb{P}(\overline{a} \cap b)$ and $\mathbb{P}(a \cap \overline{b})$ denote the posterior probability of including each variable separately. The first measure of jointness is simply $\mathbb{P}(a \cap b)$. However, this measure ignores much of the information about the relationships between the regressors. Doppelhofer & Weeks (2009) measure is defined as:

$$J_{DW} = \log\left[\frac{\mathbb{P}(a \cap b) \cdot \mathbb{P}(\overline{a} \cap \overline{b})}{\mathbb{P}(\overline{a} \cap b) \cdot \mathbb{P}(a \cap \overline{b})}\right]. \tag{38}$$

If $J_{DW} < -2$, $-2 < J_{DW} < -1$, $-1 < J_{DW} < 1$, $1 < J_{DW} < 2$, and $J_{DW} > 2$, the authors classify the regressors as strong substitutes, significant substitutes, not significantly related, significant complements, and strong complements, respectively. Jointness measure proposed by Ley & Steel (2007) is given by:

$$J_{LS} = \frac{\mathbb{P}(a \cap b)}{\mathbb{P}(\overline{a} \cap b) + \mathbb{P}(a \cap \overline{b})}. \tag{39}$$

The measure takes values in the range $[0, \infty)$, with higher values indicating a stronger complementary relationship. Finally, Hofmarcher et al. (2018) measure of jointness is:

$$J_{HCGHM} = \frac{(\mathbb{P}(a \cap b) + \rho) \cdot \mathbb{P}(\overline{a} \cap \overline{b}) + \rho) - (\mathbb{P}(\overline{a} \cap b) + \rho) \cdot \mathbb{P}(a \cap \overline{b}) + \rho)}{(\mathbb{P}(a \cap b) + \rho) \cdot \mathbb{P}(\overline{a} \cap \overline{b}) + \rho) + (\mathbb{P}(\overline{a} \cap b) + \rho) \cdot \mathbb{P}(a \cap \overline{b}) + \rho) + \rho}. \tag{40}$$

Hofmarcher et al. (2018) advocate the use of the Jeffreys (1946) prior, which results in $\rho = \frac{1}{2}$. The measure takes values from -1 to 1, where values close to -1 indicate substitutes, and those close to 1 complements.

[6] For a thorough discussion of model priors see Sala-I-Martin et al. (2004); Ley & Steel (2009); George (2010); Eicher et al. (2011).

[7] To learn more about jointness measures, we recommend reading Doppelhofer & Weeks (2009); Ley & Steel (2007); Hofmarcher et al. (2018) in that order.

# 3 Data preparation

This section demonstrates how to prepare the data for estimation. The first step involves installing the package and subsequently loading it into the R session.

```
> install.packages("bdsm")
```

```
> library(bdsm)
```

Throughout the manuscript, we use the data from Moral-Benito (2016) on the determinants of economic growth. The package includes the data along with a detailed description of all variables.

```
> ?economic_growth
```

The data used for estimation must be in a specific format. The first two columns should specify time and the entity (e.g., country). The dependent variable should be placed in the third column, while the regressors should occupy the remaining columns. The data should be arranged as follows:

```
> economic_growth[1:12,1:10]
```

```
# A tibble: 12 × 10
    year country   gdp    ish    sed    pgrw   pop   ipr    opem    gsh
   <dbl>   <dbl> <dbl>  <dbl>  <dbl>   <dbl> <dbl> <dbl>   <dbl>  <dbl>
 1  1960       1  8.25 NA     NA     NA       NA    NA   NA       NA
 2  1970       1  8.37  0.122  0.139  0.0235  10.9  61.1  1.08    0.191
 3  1980       1  8.54  0.207  0.141  0.0300  13.9  92.3  1.06    0.203
 4  1990       1  8.63  0.203  0.28   0.0303  18.9 100.   0.898   0.232
 5  2000       1  8.66  0.115  0.774  0.0215  25.3  81.2  0.636   0.219
 6  1960       2  8.97 NA     NA     NA       NA    NA   NA       NA
 7  1970       2  9.19  0.164  0.604  0.0152  20.6 103.   0.0823  0.184
 8  1980       2  9.30  0.185  0.792  0.0167  24.0 112.   0.0786  0.164
 9  1990       2  9.01  0.145  1.09   0.0154  28.4  73.8  0.104   0.174
10  2000       2  9.34  0.148  1.57   0.0130  33.0  82.6  0.180   0.174
11  1960       3  9.29 NA     NA     NA       NA    NA   NA       NA
12  1970       3  9.60  0.258  2.60   0.0219  10.3  87.4  0.215   0.143
```

However, it is common for researchers to store their data in alternative format:

```
> original_economic_growth[1:12,1:10]
```

```
# A tibble: 12 × 10
   country  year   gdp lag_gdp   ish   sed   pgrw   pop   ipr   opem
     <dbl> <dbl> <dbl>   <dbl> <dbl> <dbl>  <dbl> <dbl> <dbl>  <dbl>
 1       1  1970  8.37    8.25 0.122 0.139 0.0235  10.9  61.1 1.08
 2       1  1980  8.54    8.37 0.207 0.141 0.0300  13.9  92.3 1.06
 3       1  1990  8.63    8.54 0.203 0.28  0.0303  18.9 100.  0.898
 4       1  2000  8.66    8.63 0.115 0.774 0.0215  25.3  81.2 0.636
 5       2  1970  9.19    8.97 0.164 0.604 0.0152  20.6 103.  0.0823
 6       2  1980  9.30    9.19 0.185 0.792 0.0167  24.0 112.  0.0786
 7       2  1990  9.01    9.30 0.145 1.09  0.0154  28.4  73.8 0.104
 8       2  2000  9.34    9.01 0.148 1.57  0.0130  33.0  82.6 0.180
 9       3  1970  9.60    9.29 0.258 2.60  0.0219  10.3  87.4 0.215
10       3  1980  9.77    9.60 0.236 2.94  0.0143  12.7 119.  0.233
11       3  1990  9.92    9.77 0.238 2.90  0.0142  14.6 106.  0.266
12       3  2000 10.2     9.92 0.234 3     0.0125  16.9  95.6 0.380
```

In this case, the user can use the `join_lagged_col` function to transform the dataset into the desired format. The user needs to specify the dependent variable column (`col`), the lagged dependent variable column (`col_lagged`), the column identifying the cross-section (`entity_col`), the column with the time index (`timestamp_col`), and the change in the number of time units from period to period (`timestep`).

```
> economic_growth <- join_lagged_col(df = original_economic_growth,
+                           col = gdp, col_lagged = lag_gdp,
+                           timestamp_col = year,
+                           entity_col = country, timestep = 10)
```

Once the data is in the correct format, the user can perform further data preparation using the `feature_standardization` function. It allows to perform demeaning (entity/time effects) or scaling (standardization) as needed. Often there are columns to which the transformation should not be applied. These can be specified with the `excluded_cols`. It is also possible to group elements of the data frame with respect to a given column with the `gropu_by_col`. Finally, with the `scale` parameter we can decide whether we want to apply both demeaning and scaling or just demeaning.

For example, we can first standardize all features:

```
> data_standardized_features <- feature_standardization(
+     df            = economic_growth,
+     excluded_cols = c(country, year, gdp)
+   )
```

and then apply cross-sectional demeaning (fixed time effects):

```
> data_prepared <- feature_standardization(
+     df            = data_standardized_features,
+     group_by_col  = year,
+     excluded_cols = country,
+     scale         = FALSE
+   )
```

Note that the example below is the data preparation scheme which was used in Moral-Benito (2016). There is no need to apply panel demeaning (entity fixed effects) in this framework as can be seen in Equation 12.[8].

# 4 Estimation of the model space

The prepared data can be used to find MLEs discribed in subsection 2.2. This is done using the core function of the package: `bma_prep`. The function creates a list with two tables. We refer to the first as the the model space, because it contains the estimated MLEs of the parameters for each considered model. The second contains the values of the likelihood function at the estimated MLE, the Bayesian information criterion, and the standard errors of the parameters of interest for each model. In both tables a single column represents a single considered model.

The MLEs for the parameters are found through numerical optimization. More advanced users can use the `control` parameter to control the way the numerical optimization is performed. We refer to the function manual for more details and `stats` package for more details.

Two types of standard errors are provided, both derived from the Hessian of the maximized log-likelihood function. The first type consists of the regular standard errors, calculated using the inverse of the observed information matrix:

$$I(\hat{\theta}) = -\frac{\partial^2 l(\hat{\theta})}{\partial\theta\partial\theta'} \tag{41}$$

where $\hat{\theta}$ are the estimated MLE parameters, $I(\hat{\theta})$ is the information matrix and $l(\hat{\theta}) = \log L(\hat{\theta})$ is the natural logarithm of the likelihood function. The variance covariance matrix is given by:

[8] In theory the results should be the same with entity fixed effects. However, because we use numerical methods some changes might occur

$$Var(\hat{\theta}) = I(\hat{\theta})^{-1}, \tag{42}$$

and the standard errors by

$$\text{SE}(\hat{\theta}) = \sqrt{\text{diag}(Var(\hat{\theta}))}. \tag{43}$$

where the square root is obviously applied separately to each coordinate of the vector with diagonal values. The second type are the robust standard errors or heteroscedasticity consistent standard errors. To understand how they work, we first have to rewrite the equation Equation 17 in a form which will display the contribution of each entity on the likelihood value. First note that:

$$L(\theta) \propto -\frac{1}{2} tr\{\Sigma_{11}^{-1} U_1' U_1\} - \sum_{i=1}^{N} \frac{1}{2} \log\det \Sigma_{11} - \frac{1}{2} \log\det(\frac{H}{N}) \tag{44}$$

Now, because of the cyclic property of the trace we can rewrite the first term as:

$$-\frac{1}{2} tr\{\Sigma_{11}^{-1} U_1' U_1\} = -\frac{1}{2} tr\{U_1 \Sigma_{11}^{-1} U_1'\} = -\frac{1}{2} \sum_{i=1}^{N} u_i \Sigma_{11}^{-1} u_i' \tag{45}$$

where $u_i$ is a row vector corresponding to the data relating to the single entity $i$. Hence, the entire likelihood function can be rewritten as a sum of contributions from each entity:

$$L(\theta) \propto \sum_{i=1}^{N} -\frac{1}{2} (\log\det \Sigma_{11} + \log\det(\frac{H}{N}) + u_i \Sigma_{11}^{-1} u_i') \tag{46}$$

From there we can see that the contribution of a single entity $i$ is:

$$l_i(\theta) \propto -\frac{1}{2} (\log\det \Sigma_{11} + \log\det(\frac{H}{N}) + u_i \Sigma_{11}^{-1} u_i'). \tag{47}$$

Now if we consider a multivariate function $\mathbf{l}(\theta)$ of all such contributions (with single contributions as its coordinates), we can find it's gradient at the MLE: $G(\hat{\theta}) = \frac{\mathbf{l}(\hat{\theta})}{\partial \theta}$. Then the robust variance is:

$$Var_R(\hat{\theta}) = I(\hat{\theta})^{-1} \cdot G'(\hat{\theta}) G(\hat{\theta}) \cdot I(\hat{\theta})^{-1} \tag{48}$$

and the robust standard errors are given by:

$$\text{SE}_R(\hat{\theta}) = \sqrt{\text{diag}(\hat{V}_R(\hat{\theta}))}. \tag{49}$$

where the square root is again applied to each coordinate separately.

The `bma_prep` function is the most computationally intensive part of the package. Therefore, the function provides an option for parallel computing. If the user's dataset contains only a few regressors, the sufficient option is

```
> for_bma <- bma_prep(df = data_prepared, dep_var_col = gdp,
+                     timestamp_col = year, entity_col = country,
+                     init_value = 0.5)
```

However, for larger datasets, it is better to take advantage of parallel computing. Then the numerical optimization used to find MLEs can be computed on separate threads for each model. To do this, first load the `parallel` package and set up a cluster.

```
> library(parallel)
> # Here we try to use all available cores on the system.
> # You might want to lower the number of cores depending on your needs.
> cores <- detectCores()
> cl <- makeCluster(cores)
> setDefaultCluster(cl)
```

Then the user just needs to provide this cluster to the function:

```
> for_bma <- bma_prep(df = data_prepared, dep_var_col = gdp,
+                     timestamp_col = year, entity_col = country,
+                     init_value = 0.5, cl = cl)
```

Even with parallel computing, executing `bma_prep` may be time-consuming. Hence, the function displays a progress bar, so that the user can easily track the ongoing computation. Moreover, for users who want to practice using the `bdsm` package, we provide the already computed objects included with the package:

```
> ?bma_prep_objects_full
```

# 5 Performing Bayesian model averaging

## 5.1 Bayesian model averaging: The `bma` function

The `bma` function enables users to perform Bayesian model averaging using the object obtained with the `bma_prep` function. The `app` parameter specifies the decimal place to which the BMA statistics should be rounded in the results.

```
> bma_results <- bma(bma_prep_objects_full, df = data_prepared, round = 3)
```

The `bma` function returns a list containing 16 elements. However, most of these elements are only required for other functions. The main objects of interest are the two tables with the BMA statistics. The results obtained with binomial model prior are first on the list.

```
> bma_results[[1]]
          PIP     PM   PSD  PSDR  PMcon PSDcon PSDRcon    %(+)
gdp_lag    NA  0.918 0.075 0.107  0.918  0.075   0.107 100.000
ish     0.773  0.063 0.045 0.062  0.081  0.034   0.059 100.000
sed     0.717  0.031 0.057 0.071  0.043  0.063   0.081  70.312
pgrw    0.714  0.018 0.030 0.052  0.025  0.033   0.060  99.609
pop     0.990  0.121 0.062 0.079  0.122  0.061   0.079 100.000
ipr     0.657 -0.033 0.032 0.043 -0.050  0.027   0.044   0.000
opem    0.766  0.034 0.029 0.032  0.044  0.026   0.030 100.000
gsh     0.751 -0.013 0.039 0.086 -0.017  0.044   0.099  28.906
lnlex   0.864  0.086 0.072 0.095  0.100  0.069   0.095 100.000
polity  0.678 -0.056 0.046 0.052 -0.083  0.030   0.043   0.000
```

PIP denotes the posterior inclusion probability, PM denotes the posterior mean, PSD denotes the posterior standard deviation, and PSDR denotes the posterior standard deviation calculated using robust standard errors. These are the four main results of BMA with respect to the assessment of individual regressors. PMcon, PSDcon, and PSDRcon denote the posterior mean, posterior standard deviation, and posterior standard deviation based on robust standard errors, respectively, conditional on the inclusion of the variable. Users should base their interpretation of the results on conditional BMA statistics only when they believe that certain regressors must be included. Finally, for a given parameter we can consider all models that include this parameter, and check if it has a positive or negative value. %(+) denotes the percentage of models with positive value for a given parameter across all models that include that parameter. A value of %(+) equal to 0% or 100% indicates coefficient sign stability.

The PIP for all the regressors shows that none of them can be considered very strong according to the classification by Raftery (1995). This also applies to the population variable (`pop`), which has a PIP of 0.990 due solely to approximation. These results are corroborated by the ratios of PM to PSD and PSDR. In particular, for the absolute value of the PM to PSDR ratio, only the population variable exceeds 1.3, while investment (`ish`) and the democracy index (`polity`) are above 1. This finding led Moral-Benito

(2016) to emphasize the fragility of economic growth determinants. The only variable that can be considered robust across all metrics is the lagged GDP (`gdp_lag`). However, the results change when using the binomial-beta model prior, which is included as the second object in the `bma` list.

```
> bma_results[[2]]
          PIP     PM   PSD  PSDR  PMcon PSDcon PSDRcon    %(+)
gdp_lag    NA  0.922 0.075 0.112  0.922  0.075   0.112 100.000
ish     0.954  0.074 0.034 0.060  0.078  0.031   0.059 100.000
sed     0.939  0.048 0.057 0.070  0.051  0.057   0.071  70.312
pgrw    0.939  0.024 0.032 0.057  0.025  0.032   0.059  99.609
pop     0.998  0.101 0.057 0.074  0.101  0.057   0.074 100.000
ipr     0.924 -0.044 0.028 0.040 -0.048  0.025   0.040   0.000
opem    0.953  0.036 0.024 0.026  0.038  0.024   0.026 100.000
gsh     0.948 -0.018 0.041 0.088 -0.019  0.042   0.090  28.906
lnlex   0.974  0.115 0.063 0.089  0.118  0.061   0.088 100.000
polity  0.929 -0.077 0.036 0.046 -0.083  0.030   0.042   0.000
```

In the case of the binomial-beta model prior, the PIPs for all the regressors increase. Population is classified as very strong, while all other regressors are classified as strong or positive according to posterior inclusion probabilities. There are also considerable changes in the PM to PSD and PSD_R ratios. The absolute value of the PM to PSD ratio exceeds two for investment and the democracy index, and is above 1.3 for population, investment price (`ipr`), trade openness (`open`), and life expectancy (`lnlex`). However, these results are less pronounced when using robust standard errors, with only population, trade openness, and the democracy index remaining above 1.3. Consequently, the results are not robust with respect to the choice of prior model specification. The reasons behind these differences will become clear once other functionalities of the package are explored.

The last object in the list is a table containing the prior and posterior expected model sizes for the binomial and binomial-beta model priors. Importantly, these numbers reflect only the number of regressors in a model and do not include the lagged dependent variable, which is present in every model by construction.

```
> bma_results[[16]]
              Prior models size Posterior model size
Binomial                    4.5                6.910
Binomial-beta               4.5                8.558
```

The results show that, after observing the data, the researcher should expect around seven and eight and a half regressors in the model under the binomial and binomial-beta model priors, respectively. These numbers may seem high; however, they are driven by relatively substantial PIPs. This illustrates the importance of focusing on both posterior inclusion probabilities and the ratios of posterior mean to posterior standard deviation when assessing the robustness of the regressors.

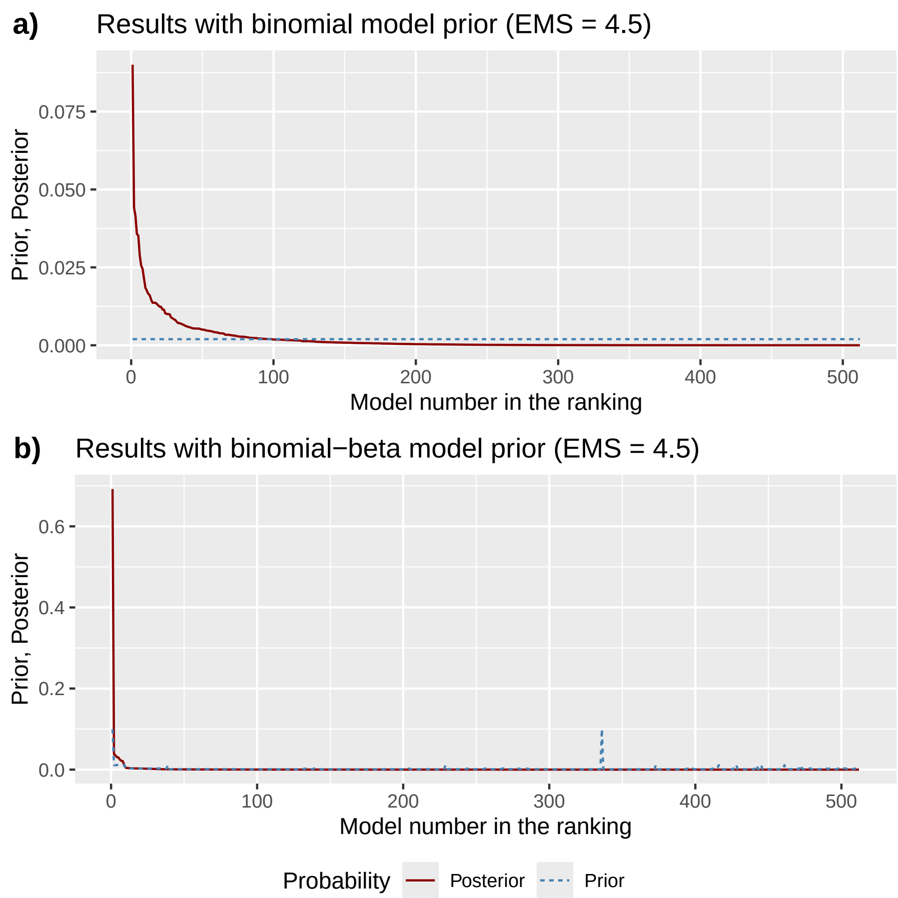
## 5.2 Prior and posterior model probabilities

The `model_pmp` function allows the user to compare prior and posterior model probabilities over the entire model space in the form of a graph. The models are ranked from the one with the highest to the one with the lowest posterior model probability. The function returns a list with three objects:

1. a graph for the binomial model prior;
2. a graph for the binomial-beta model prior;
3. a combined graph for both binomial and binomial-beta model priors.

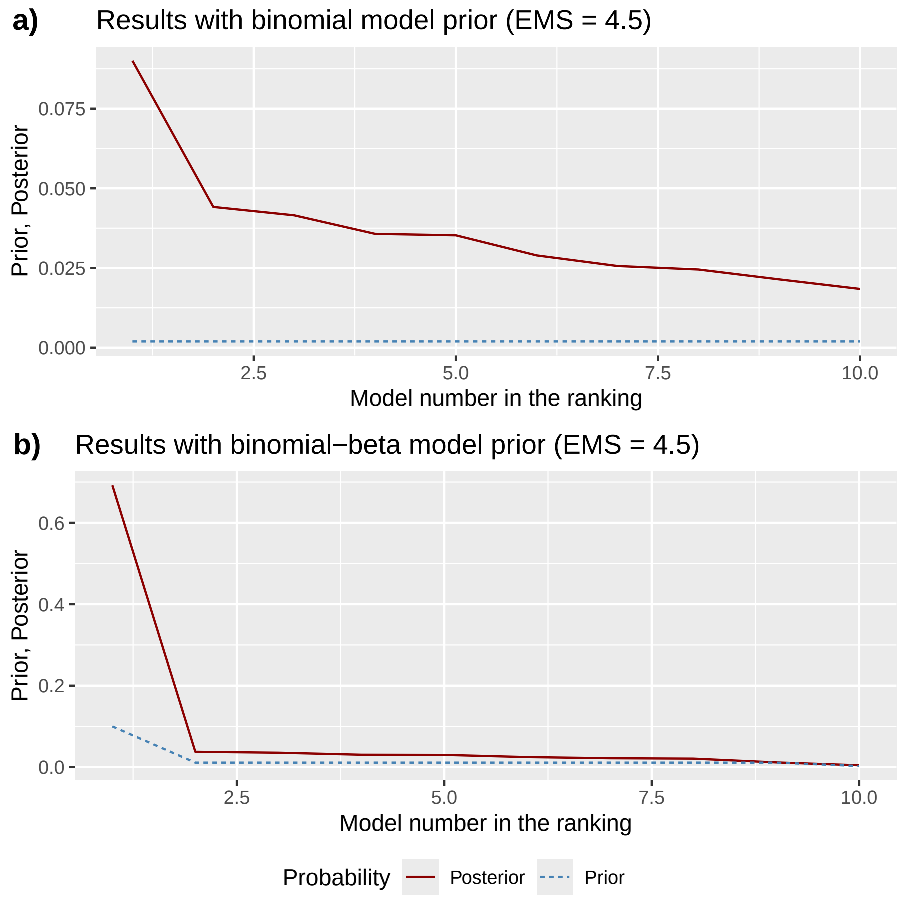
The user can retrieve each graph separately from the list; however, the function automatically displays a combined graph.

```
> for_models <- model_pmp(bma_results)
```




The graphs demonstrate that most of the posterior probability mass is concentrated within just a couple of models. To view the results for only the best models, the user can use the `top` parameter.

```
> for_models <- model_pmp(bma_results, top = 10)
```



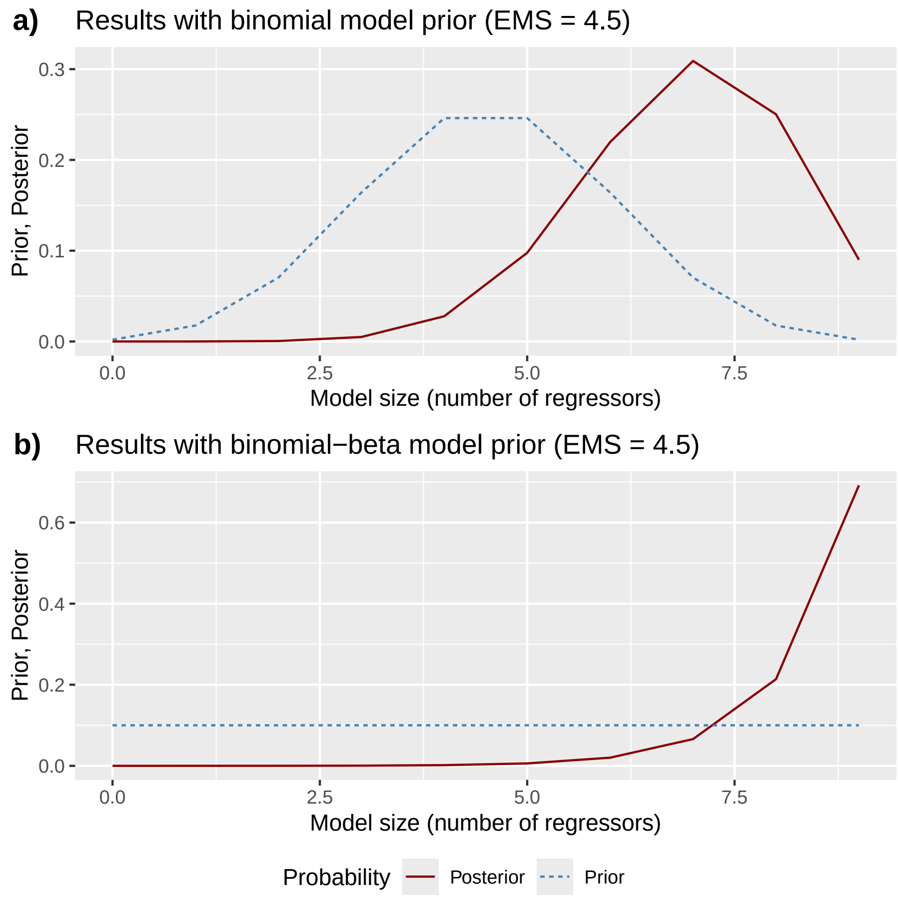

The last graph for the binomial-beta prior is particularly illuminating in terms of explaining the very high values of posterior inclusion probabilities. Almost 70% of the posterior probability mass is concentrated in just one model; therefore, variables included in this model will have very high PIP values. The model in question will be identified after implementing `model_sizes` (and `best_models`, which is covered in subsection 5.3). Nevertheless, the results from the graph suggest that the best model is the one that includes all the regressors or none (because the prior value is around $\frac{1}{9}$ on the plot).

The `model_sizes` function displays prior and posterior model probabilities on a graph for models of different sizes. The graphs exclude the lagged dependent variable; therefore, the model with zero regressors still includes the lagged dependent variable. Similarly to the `model_pmp` function is returns a list with three objects:

1. a graph for the binomial model prior;
2. a graph for the binomial-beta model prior;
3. a combined graph for both binomial and binomial-beta model priors.

Again, the user can retrieve each graph separately from the list; however, the function automatically displays a combined graph.

```
> size_graphs <- model_sizes(bma_results)
```




The graph in panel b) again explains why PIPs are so high in the case of the binomial-beta model prior. The model with all the regressors accounts for almost 70% of the total posterior probability mass, while the remaining portion is concentrated on models with a high number of regressors. In contrast, the posterior probability mass for the binomial model prior is centered around models with seven regressors. This graph clearly illustrates the impact of changes in the model prior on posterior probabilities.

## 5.3 Selecting the best models

The `best_models` function allows the user to view a chosen number of the best models in terms of posterior model probability. The function returns a list containing nine objects:

1. An inclusion table stored as an array object;
2. A table with estimation results using regular standard errors, stored as an array object;
3. A table with estimation results using robust standard errors, stored as an array object;
4. An inclusion table stored as a knitr object;
5. A table with estimation results using regular standard errors, stored as a knitr object;
6. A table with estimation results using robust standard errors, stored as a knitr object;
7. An inclusion table stored as a gTree object;
8. A table with estimation results using regular standard errors, stored as a gTree object;

9. A table with estimation results using robust standard errors, stored as a gTree object;

The parameters `estimate` and `robust` pertain only to the results that will be automatically displayed after running the function. The parameter `criterion` determines whether the models should be ranked according to posterior model probabilities calculated using the binomial (1) or binomial-beta (2) model prior. To obtain the inclusion array for the 10 best models ranked with the binomial model prior, the user needs to run:

```
> best_8_models <- best_models(bma_results, criterion = 1, best = 8)
> best_8_models[[1]]
```

```
        'No. 1' 'No. 2' 'No. 3' 'No. 4' 'No. 5' 'No. 6' 'No. 7' 'No. 8'
gdp_lag    1.00   1.000   1.000   1.000   1.000   1.000   1.000   1.000
ish        1.00   1.000   1.000   1.000   1.000   1.000   1.000   0.000
sed        1.00   1.000   1.000   0.000   1.000   1.000   1.000   1.000
pgrw       1.00   1.000   1.000   1.000   0.000   1.000   1.000   1.000
pop        1.00   1.000   1.000   1.000   1.000   1.000   1.000   1.000
ipr        1.00   0.000   1.000   1.000   1.000   1.000   1.000   1.000
opem       1.00   1.000   1.000   1.000   1.000   1.000   0.000   1.000
gsh        1.00   1.000   1.000   1.000   1.000   0.000   1.000   1.000
lnlex      1.00   1.000   1.000   1.000   1.000   1.000   1.000   1.000
polity     1.00   1.000   0.000   1.000   1.000   1.000   1.000   1.000
PMP        0.09   0.044   0.042   0.036   0.035   0.029   0.026   0.025
```

1 indicates the presence of a given regressor in a model, while the last row displays the posterior model probability of that model. To obtain a knitr table with estimation output with regular standard errors for best 3 models ranked with binomial-beta model prior, the user needs to run:

```
> best_3_models <- best_models(bma_results, criterion = 2, best = 3)
> best_3_models[[5]]
```

```
|        |      'No. 1'       |       'No. 2'       |       'No. 3'       |
|:-------|:------------------:|:-------------------:|:-------------------:|
|gdp_lag | 0.924 (0.075)***   | 0.892 (0.073)***    |  0.92 (0.055)***    |
|ish     |  0.077 (0.03)**    | 0.092 (0.029)***    |  0.055 (0.029)*     |
|sed     |   0.053 (0.055)    |   0.007 (0.059)     |  0.081 (0.049)*     |
|pgrw    |   0.025 (0.032)    |   0.017 (0.032)     |   0.037 (0.032)     |
|pop     |  0.095 (0.055)*    | 0.139 (0.053)***    |  0.106 (0.053)**    |
|ipr     |  -0.047 (0.025)*   |         NA          | -0.061 (0.023)***   |
|opem    |   0.036 (0.023)    |   0.032 (0.023)     |  0.041 (0.023)*     |
|gsh     |  -0.019 (0.041)    |   -0.024 (0.04)     |  -0.022 (0.045)     |
|lnlex   |  0.124 (0.058)**   |   0.054 (0.058)     |  0.132 (0.054)**    |
|polity  | -0.083 (0.029)***  | -0.093 (0.029)***   |         NA          |
|PMP     |       0.692        |       0.038         |       0.035         |
```

The comparison of the last two tables further highlights the importance of the model prior. The best model under the binomial model prior accounts for around 9% of the posterior probability mass, while the best model under the binomial-beta model prior accounts for over 69%. Finally, to obtain a gTree table with estimation output using robust standard errors for the top 3 models ranked by the binomial-beta model prior, the user needs to run:

```
> best_3_models <- best_models(bma_results, criterion = 2, best = 3)
> best_3_models[[9]]
```

```
gTree[GRID.gTree.2013]
```

| | 'No. 1' | 'No. 2' | 'No. 3' |
|---|---|---|---|
| *gdp_lag* | 0.924 (0.112)*** | 0.892 (0.097)*** | 0.92 (0.089)*** |
| *ish* | 0.077 (0.059) | 0.092 (0.055)*** | 0.055 (0.052) |
| *sed* | 0.053 (0.068) | 0.007 (0.079) | 0.081 (0.071) |
| *pgrw* | 0.025 (0.058) | 0.017 (0.058) | 0.037 (0.06) |
| *pop* | 0.095 (0.072) | 0.139 (0.07)*** | 0.106 (0.073) |
| *ipr* | −0.047 (0.039) | NA | −0.061 (0.045)*** |
| *opem* | 0.036 (0.024) | 0.032 (0.025) | 0.041 (0.03) |
| *gsh* | −0.019 (0.087) | −0.024 (0.081) | −0.022 (0.109) |
| *lnlex* | 0.124 (0.085) | 0.054 (0.071) | 0.132 (0.098) |
| *polity* | −0.083 (0.042)*** | −0.093 (0.042)*** | NA |
| *PMP* | 0.692 | 0.038 | 0.035 |

The comparison of the last two tables and the estimation outputs with regular and robust standard errors demonstrates how the results change when switching between these two variance estimators.

## 5.4 Calculating jointness measures

Within the BMA framework, it is possible to establish the nature of the relationship between pairs of examined regressors using the jointness measures. This can be accomplished using the `jointness` function. The latest jointness measure, introduced by Hofmarcher et al. (2018), has been shown to outperform older alternatives developed by Ley & Steel (2007) and Doppelhofer & Weeks (2009)[9]. Therefore, the Hofmarcher et al. (2018) measure is the default option in the `jointness` function.

```
> jointness(bma_results)
```

```
         ish   sed  pgrw   pop   ipr  opem   gsh lnlex polity
ish       NA 0.217 0.208 0.531 0.151 0.262 0.244 0.366  0.182
sed    0.806    NA 0.155 0.421 0.116 0.200 0.189 0.289  0.126
pgrw   0.806 0.779    NA 0.416 0.124 0.199 0.187 0.284  0.132
pop    0.906 0.874 0.874    NA 0.304 0.517 0.490 0.711  0.346
ipr    0.782 0.758 0.759 0.846    NA 0.153 0.139 0.210  0.102
opem   0.830 0.802 0.803 0.902 0.781    NA 0.241 0.373  0.170
gsh    0.822 0.795 0.795 0.894 0.773 0.820    NA 0.341  0.154
lnlex  0.865 0.836 0.836 0.944 0.811 0.863 0.854    NA  0.227
polity 0.791 0.764 0.765 0.856 0.745 0.788 0.780 0.818     NA
```

Above the main diagonal the user can find the results for the binomial model prior, and below the results for the binomial-beta model prior. All the values in the table are

[9]See section 2.4 for the interpretations of jointness measures.

positive, indicating complementary relationships between the regressors. Notably, the values for the binomial-beta prior are substantially higher than those for the binomial prior. This result is not surprising, as the model with all the regressors accounts for almost 70% of the total posterior probability mass.

To obtain the results for the Ley & Steel (2007) measure, the user should run:

```
> jointness(bma_results, measure = "LS")
```

```
          ish    sed   pgrw    pop   ipr   opem    gsh  lnlex polity
ish        NA  1.470  1.448  3.285 1.230  1.678  1.603  2.213  1.324
sed     9.551     NA  1.250  2.467 1.091  1.424  1.378  1.812  1.141
pgrw    9.539  8.284     NA  2.437 1.100  1.416  1.369  1.792  1.146
pop    20.258 14.941 14.953     NA 1.876  3.162  2.935  5.986  2.062
ipr     8.387  7.441  7.492 12.022    NA  1.223  1.178  1.477  1.014
opem   11.040  9.361  9.375 19.533 8.307     NA  1.583  2.212  1.292
gsh    10.492  8.999  9.012 17.822 7.998 10.341     NA  2.061  1.241
lnlex  14.094 11.424 11.425 35.099 9.745 13.898 12.960     NA  1.566
polity  8.777  7.686  7.724 12.875 7.014  8.630  8.293 10.198     NA
```

The values corroborate the results obtained using the Hofmarcher et al. (2018) measure. All the regressors exhibit complementary relationships, which are visibly stronger under the binomial-beta model prior.

However, the Doppelhofer & Weeks (2009) measure yields a slightly different outcome:

```
> jointness(bma_results, measure = "DW")
```

```
         ish   sed   pgrw    pop    ipr   opem   gsh  lnlex polity
ish       NA 0.051  0.020  0.005  0.018  0.030 0.008 -0.004  0.067
sed    0.989    NA -0.023 -0.029  0.004 -0.008 0.000  0.005 -0.024
pgrw   0.974 0.905     NA -0.001  0.047 -0.001 0.001 -0.007  0.010
pop    1.019 0.957  0.987     NA -0.018 -0.023 0.049  0.154  0.012
ipr    0.985 0.931  0.980  0.974     NA  0.048 0.019  0.025  0.036
opem   1.012 0.938  0.949  0.991  1.006     NA 0.032  0.139  0.032
gsh    0.972 0.931  0.941  1.042  0.962  0.991    NA  0.034  0.001
lnlex  0.983 0.957  0.954  1.173  0.985  1.105 1.001     NA -0.056
polity 1.013 0.903  0.939  0.994  0.967  0.980 0.934  0.906     NA
```

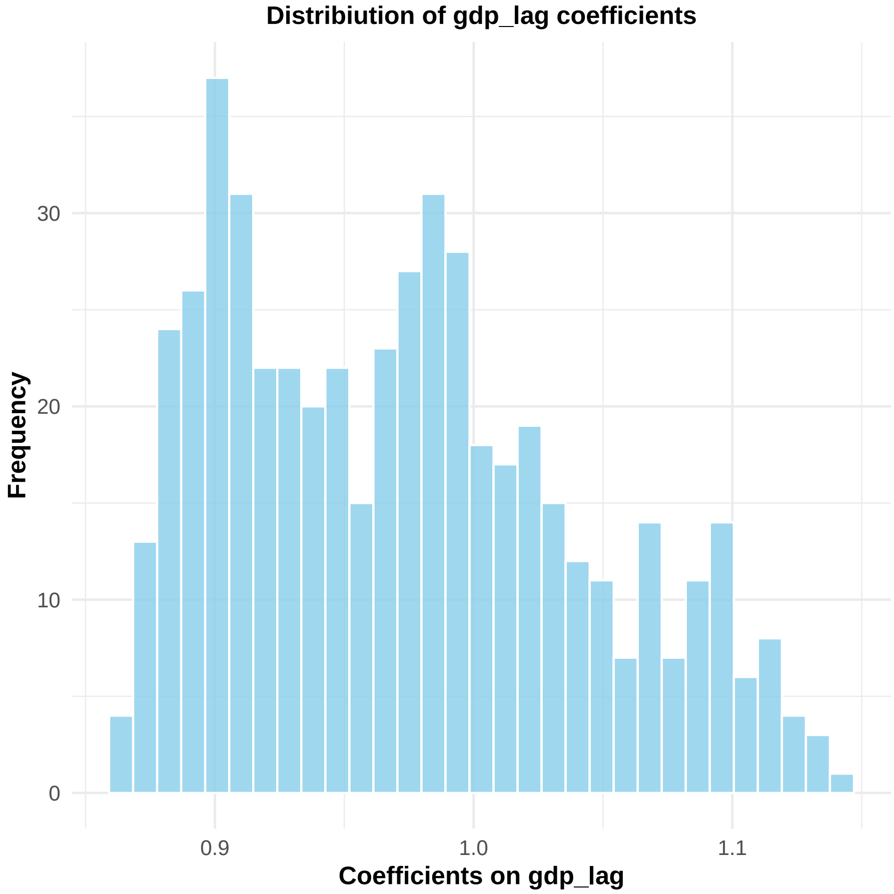
In this case, some pairs of regressors have negative values of the jointness measure under the binomial model prior; however, these values are very close to zero, indicating unrelated variables. Once again, the values for the binomial-beta model prior are higher, demonstrating how the results are influenced by the choice of model prior.

## 5.5 Visualizing model coefficients

The `coef_hist` function allows the user to plot the distribution of estimated coefficients. It returns a list containing a number of objects equal to the number of regressors plus one. The first object in the list is a graph of the coefficients for the lagged dependent variable, while the remaining objects are graphs of the coefficients for the other regressors. The graph for the lagged dependent variable collects coefficients from the entire model space, whereas the graphs for the other regressors only collect coefficients from the models that include the given regressor (half of the model space).

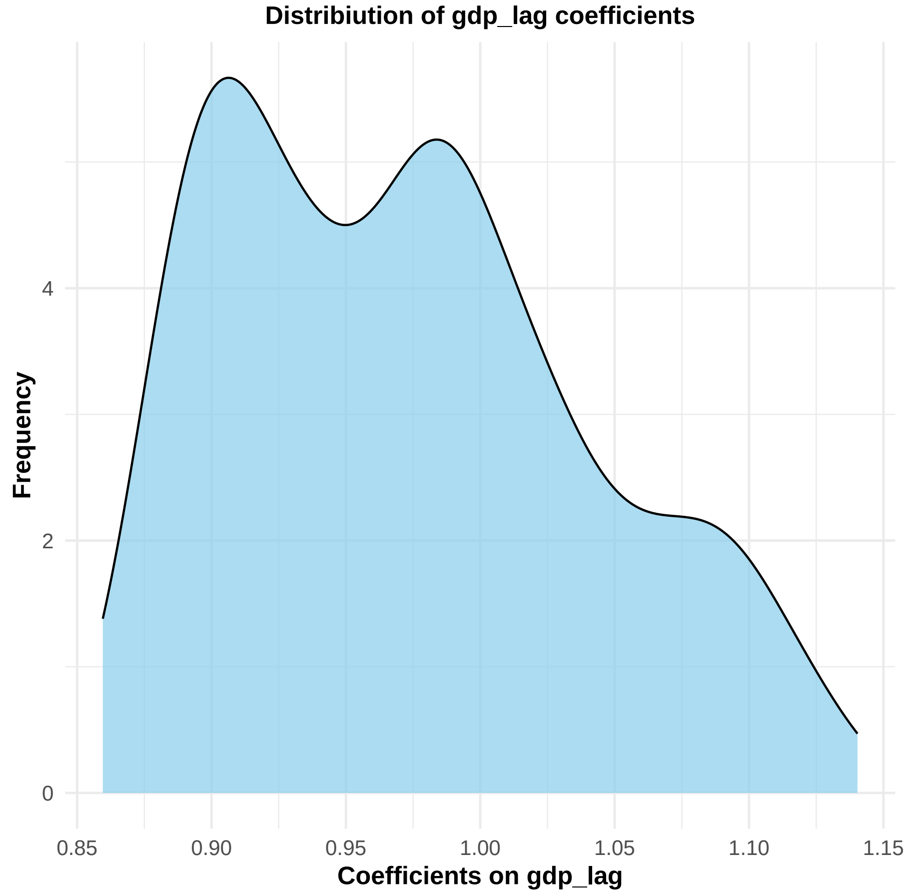
There are two main options for visualizing the coefficient distributions. The first option uses a histogram. The `coef_hist` function provides the user with options for controlling the bin widths of the histogram (`BW`, `binW`, `BN`, and `num`). The default is `BW = FD`, which selects bin widths using the Freedman-Diaconis method.

```
> coef_plots <- coef_hist(bma_results)
> coef_plots[[1]]
```



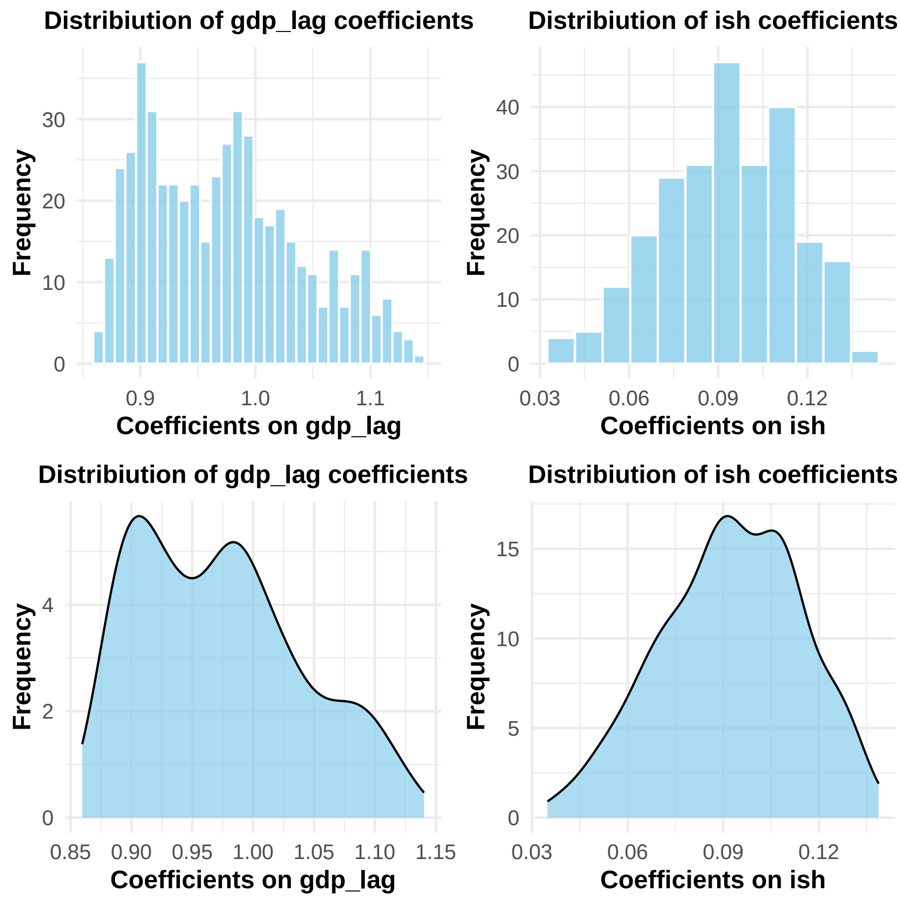

The second option allows the user to plot kernel densities.

```
> coef_plots2 <- coef_hist(bma_results, kernel = 1)
> coef_plots2[[1]]
```




The choice of appropriate plotting options is left to the user's preferences regarding the style of presentation and the size of the model space.

```
> library(gridExtra)
> grid.arrange(coef_plots[[1]], coef_plots[[2]], coef_plots2[[1]],
+              coef_plots2[[2]], nrow = 2, ncol = 2)
```




# 6 Changes in model priors

This section provides a more detailed description of the available model prior options. Subsection 6.1 discusses the consequences of changes in the expected model size, while subsection 6.2 describes the dilution prior.

## 6.1 Changing expected model size

The `bma` function calculates BMA statistics using both the binomial and binomial-beta model priors. By default, the `bma` function sets the expected model size (EMS) to $K/2$, where $K$ denotes the total number of regressors. The binomial model prior with $EMS = K/2$ leads to a uniform model prior, assigning equal probabilities to all models. In contrast, the binomial-beta model prior with $EMS = K/2$ assumes equal probabilities across all model sizes. However, the user can modify the prior model specification by changing the `EMS` parameter.

First, consider the consequence of concentrating prior probability mass on small models by setting $EMS = 2$.

```
> bma_results2 <- bma(bma_prep_objects_full, df = data_prepared,
+                     round = 3, EMS = 2)
```

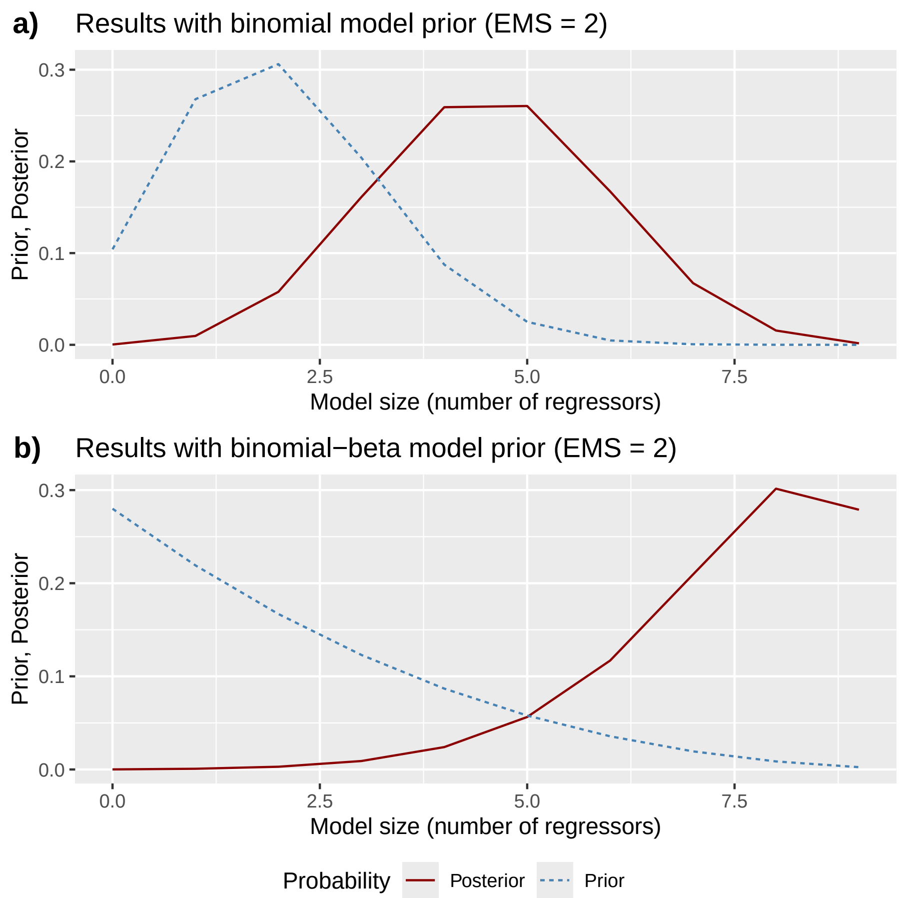
Before turning to the main BMA results, let us focus on the changes in the posterior probability mass with respect to model sizes.

```
> bma_results2[[16]]
```

```
                 Prior models size Posterior model size
Binomial                         2                4.560
Binomial-beta                    2                7.502
```

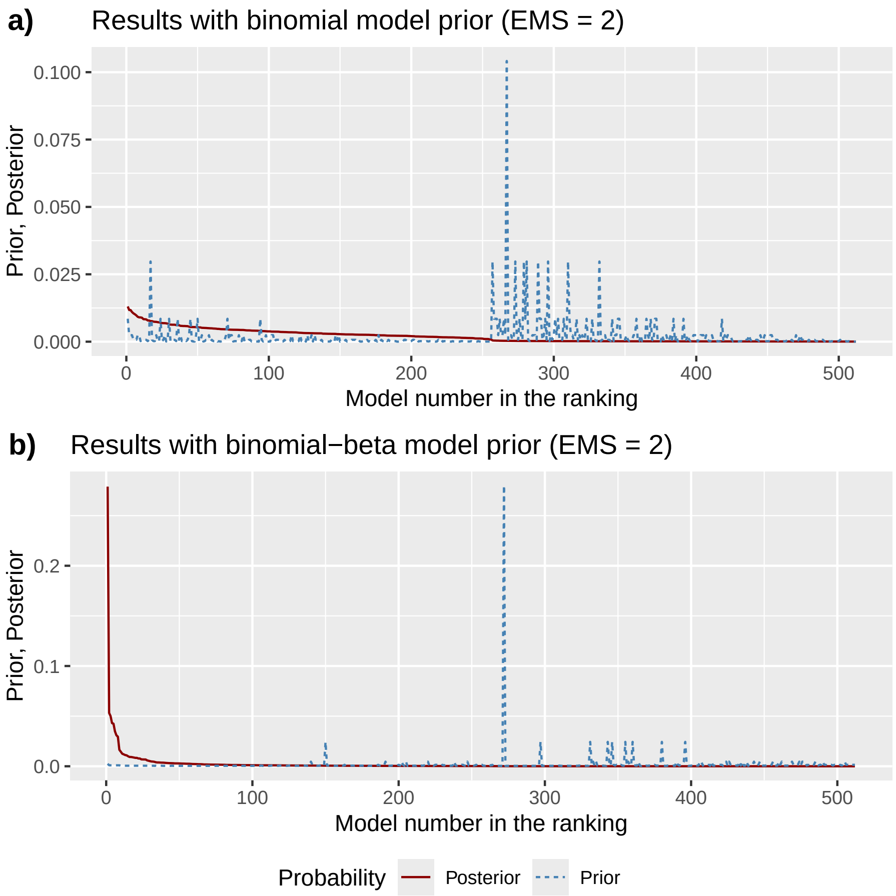
The results show that decreasing the prior expected model size led to a considerable decline in the posterior expected model size. The consequences of this change in the prior expected model size are best illustrated using the prior and posterior probability mass over model sizes.

```
> size_graphs2 <- model_sizes(bma_results2)
```




For both the binomial and binomial-beta model priors, the prior probability mass is more concentrated on small model sizes. However, for the binomial model prior, the center of the posterior probability mass shifted to medium-sized models, while it remained on large models for the binomial-beta model prior. Nevertheless, the posterior model probability for the model with all regressors decreased from nearly 0.7 for $EMS = 4.5$ to less than 0.3. There are also substantial changes in the distribution of the posterior probability mass over the model space.

```
> model_graphs2 <- model_pmp(bma_results2)
```




Both panels of the graph show that the prior and posterior model probabilities have substantially decoupled from each other. This strongly indicates that the prior and the data are suggesting vastly different model choices. The tall blue spike represents the model with no regressors. The main BMA posterior statistic for the binomial model prior also experienced a significant change.

```
> bma_results2[[1]]
```

| | PIP | PM | PSD | PSDR | PMcon | PSDcon | PSDRcon | %(+) |
|---|---|---|---|---|---|---|---|---|
| gdp_lag | NA | 0.925 | 0.080 | 0.102 | 0.925 | 0.080 | 0.102 | 100.000 |
| ish | 0.483 | 0.043 | 0.050 | 0.059 | 0.088 | 0.034 | 0.057 | 100.000 |
| sed | 0.420 | 0.014 | 0.046 | 0.058 | 0.034 | 0.065 | 0.085 | 70.312 |
| pgrw | 0.414 | 0.009 | 0.025 | 0.041 | 0.022 | 0.034 | 0.061 | 99.609 |
| pop | 0.964 | 0.142 | 0.065 | 0.082 | 0.147 | 0.060 | 0.078 | 100.000 |
| ipr | 0.344 | -0.019 | 0.031 | 0.037 | -0.056 | 0.028 | 0.045 | 0.000 |
| opem | 0.468 | 0.024 | 0.032 | 0.033 | 0.052 | 0.026 | 0.030 | 100.000 |
| gsh | 0.459 | -0.003 | 0.032 | 0.071 | -0.007 | 0.047 | 0.105 | 28.906 |
| lnlex | 0.637 | 0.052 | 0.068 | 0.087 | 0.082 | 0.069 | 0.096 | 100.000 |
| polity | 0.372 | -0.029 | 0.042 | 0.046 | -0.079 | 0.031 | 0.043 | 0.000 |

Posterior inclusion probabilities drop considerably for all the regressors, except for population, which remains almost unchanged. Interestingly, the ratios for all variables declined, with population being the exception. The ratio for population remains above two for regular standard errors and 1.7 for robust standard errors. This outcome indicates that population performs relatively better in smaller models. The results for binomial-beta model prior are given below.

```
> bma_results2[[2]]
```

| | PIP | PM | PSD | PSDR | PMcon | PSDcon | PSDRcon | %(+) |
|---|---|---|---|---|---|---|---|---|

```
gdp_lag     NA   0.919 0.075 0.109  0.919  0.075   0.109 100.000
ish      0.838   0.067 0.042 0.061  0.080  0.033   0.059 100.000
sed      0.796   0.037 0.057 0.071  0.047  0.061   0.077  70.312
pgrw     0.795   0.020 0.031 0.054  0.025  0.033   0.059  99.609
pop      0.992   0.114 0.061 0.078  0.115  0.061   0.078 100.000
ipr      0.754  -0.037 0.031 0.042 -0.049  0.026   0.042   0.000
opem     0.833   0.034 0.028 0.030  0.041  0.025   0.028 100.000
gsh      0.822  -0.015 0.040 0.087 -0.018  0.043   0.095  28.906
lnlex    0.902   0.097 0.071 0.094  0.107  0.066   0.093 100.000
polity   0.769  -0.064 0.044 0.051 -0.083  0.030   0.043   0.000
```

The change in PIPs is again significant, though not as pronounced as in the case of the binomial model prior. Changes in the ratios are relatively small and irregular for both regular and robust standard errors. The most pronounced change is the drop in the value of the ratios for the democracy index (`polity`), indicating that this regressor performs better in larger models.

It is also very instructive to examine the jointness measures calculated under the new prior specification.

```
> jointness(bma_results2, measure = "HCGHM", rho = 0.5, round = 3)
```

```
         ish   sed  pgrw    pop    ipr   opem    gsh  lnlex polity
ish       NA 0.021 0.008 -0.030  0.003  0.002  0.001 -0.012  0.021
sed    0.441    NA 0.017 -0.146  0.043  0.007  0.011 -0.036  0.026
pgrw   0.437 0.390    NA -0.155  0.053  0.012  0.011 -0.041  0.037
pop    0.667 0.586 0.583     NA -0.281 -0.057 -0.072  0.253 -0.231
ipr    0.391 0.355 0.361  0.503     NA  0.021  0.023 -0.065  0.072
opem   0.483 0.430 0.430  0.657  0.391     NA  0.010  0.022  0.020
gsh    0.467 0.419 0.418  0.636  0.378  0.464     NA -0.012  0.019
lnlex  0.559 0.497 0.495  0.793  0.439  0.562  0.538     NA -0.072
polity 0.413 0.364 0.369  0.532  0.340  0.405  0.390  0.453     NA
```

On the one hand, the results obtained with the binomial-beta model prior did not change in any significant manner. On the other hand, the results obtained with the binomial model prior changed substantially. The measure indicates that population is a substitute for both the investment price (`ipr`) and the democracy index, as well as, to a lesser extent, secondary education (`sed`) and population growth (`pgrw`).

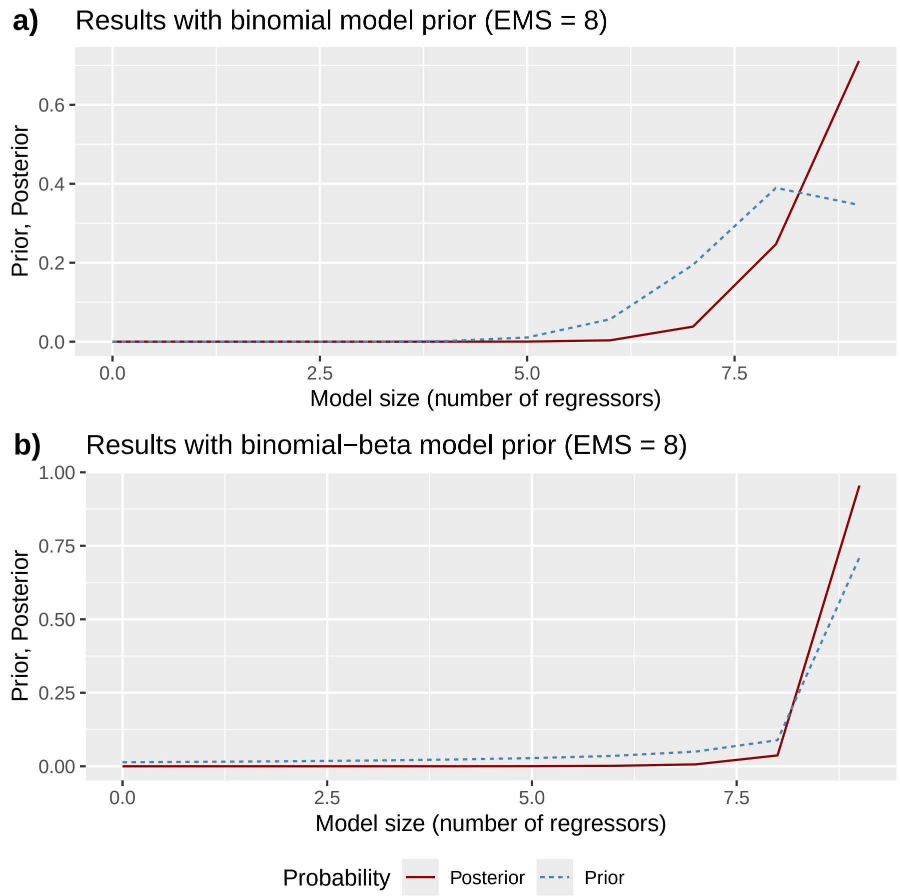
Next, to consider the consequences of concentrating prior probability mass on large models, `EMS` was set to eight.

```
> bma_results8 <- bma(bma_prep_objects_full, df = data_prepared,
+                     round = 3, EMS = 8)
> bma_results8[[16]]
```

```
              Prior models size Posterior model size
Binomial                      8                8.666
Binomial-beta                 8                8.944
```

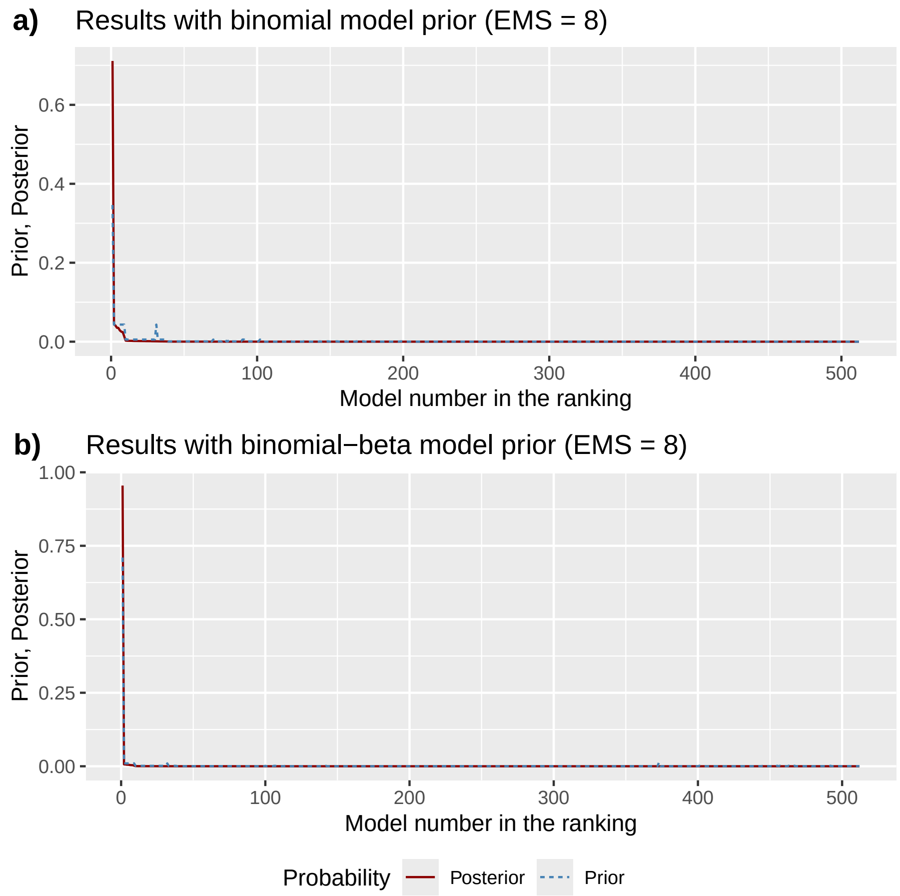
The posterior model size increased for the binomial prior; however, it remained almost unchanged for the binomial-beta model prior. The most interesting aspect is the new graphs of prior and posterior probability mass over the model sizes.

```
> size_graphs8 <- model_sizes(bma_results8)
```




In both cases, the posterior probability mass has concentrated near the models with all the regressors. However, in the case of the binomial-beta model prior, the model with all the regressors captures most of the posterior probability mass (almost 96%). This conclusion is further supported by the graphs of posterior model probability across the entire model space.

```
> model_graphs8 <- model_pmp(bma_results8)
```




Panel (a) demonstrates that the change in the expected model size led to a substantial increase in the posterior model probability for the model with all regressors under the binomial model prior. It now accounts for over 70% of the total posterior probability mass. The increase in the expected model size also influenced the main BMA statistics.

```
> bma_results8[[1]]
```

```
          PIP     PM   PSD  PSDR  PMcon PSDcon PSDRcon    %(+)
gdp_lag    NA  0.922 0.075 0.112  0.922  0.075   0.112 100.000
ish     0.967  0.075 0.034 0.060  0.077  0.031   0.059 100.000
sed     0.953  0.049 0.057 0.070  0.051  0.057   0.071  70.312
pgrw    0.953  0.024 0.032 0.057  0.025  0.032   0.058  99.609
pop     0.999  0.100 0.057 0.074  0.100  0.057   0.074 100.000
ipr     0.942 -0.045 0.027 0.040 -0.048  0.025   0.040   0.000
opem    0.965  0.036 0.024 0.026  0.037  0.024   0.025 100.000
gsh     0.961 -0.019 0.041 0.088 -0.019  0.042   0.090  28.906
lnlex   0.981  0.117 0.062 0.088  0.119  0.061   0.088 100.000
polity  0.945 -0.078 0.034 0.045 -0.083  0.030   0.043   0.000
```

The PIPs increased considerably. Population is classified as very strong, while the other regressors are classified as strong or positive. Interestingly, all the ratios have improved as well, except for population. The change in the results for the binomial-beta model prior is less pronounced.

```
> bma_results8[[2]]
```

```
          PIP     PM   PSD  PSDR  PMcon PSDcon PSDRcon    %(+)
gdp_lag    NA  0.923 0.075 0.112  0.923  0.075   0.112 100.000
ish     0.994  0.076 0.031 0.060  0.077  0.030   0.059 100.000
```

```
sed     0.992  0.052 0.055 0.068  0.053  0.055   0.068  70.312
pgrw    0.992  0.025 0.032 0.058  0.025  0.032   0.058  99.609
pop     1.000  0.096 0.055 0.072  0.096  0.055   0.072 100.000
ipr     0.990 -0.047 0.025 0.039 -0.047  0.025   0.039   0.000
opem    0.994  0.036 0.023 0.024  0.036  0.023   0.024 100.000
gsh     0.994 -0.019 0.041 0.087 -0.019  0.041   0.088  28.906
lnlex   0.997  0.123 0.058 0.086  0.123  0.058   0.086 100.000
polity  0.991 -0.082 0.030 0.043 -0.083  0.029   0.042   0.000
```

With the increase in expected model size, population is classified as very strong, and all the other regressors are classified as strong in terms of the posterior inclusion probability criterion. Similarly to the case of the binomial prior, all the ratios increased except for population.

Again, it is instructive to examine the jointness measures.

```
> jointness(bma_results8, measure = "HCGHM", rho = 0.5, round = 3)
```

```
          ish   sed  pgrw   pop   ipr  opem   gsh lnlex polity
ish        NA 0.840 0.841 0.931 0.819 0.865 0.857 0.896  0.826
sed     0.975    NA 0.814 0.903 0.792 0.837 0.830 0.869  0.799
pgrw    0.975 0.971    NA 0.904 0.793 0.838 0.830 0.870  0.800
pop     0.988 0.984 0.984    NA 0.881 0.928 0.920 0.960  0.888
ipr     0.972 0.968 0.968 0.980    NA 0.816 0.808 0.847  0.778
opem    0.978 0.974 0.974 0.988 0.971    NA 0.854 0.894  0.823
gsh     0.977 0.973 0.973 0.987 0.970 0.977    NA 0.886  0.815
lnlex   0.983 0.979 0.979 0.993 0.976 0.983 0.982    NA  0.854
polity  0.973 0.969 0.969 0.982 0.966 0.972 0.971 0.977     NA
```

The values of the measures show that all the regressors exhibit a very strong complementary relationship. This outcome, once again, underscores the importance of carefully considering the prior when interpreting jointness measures.

## 6.2 Dilution prior

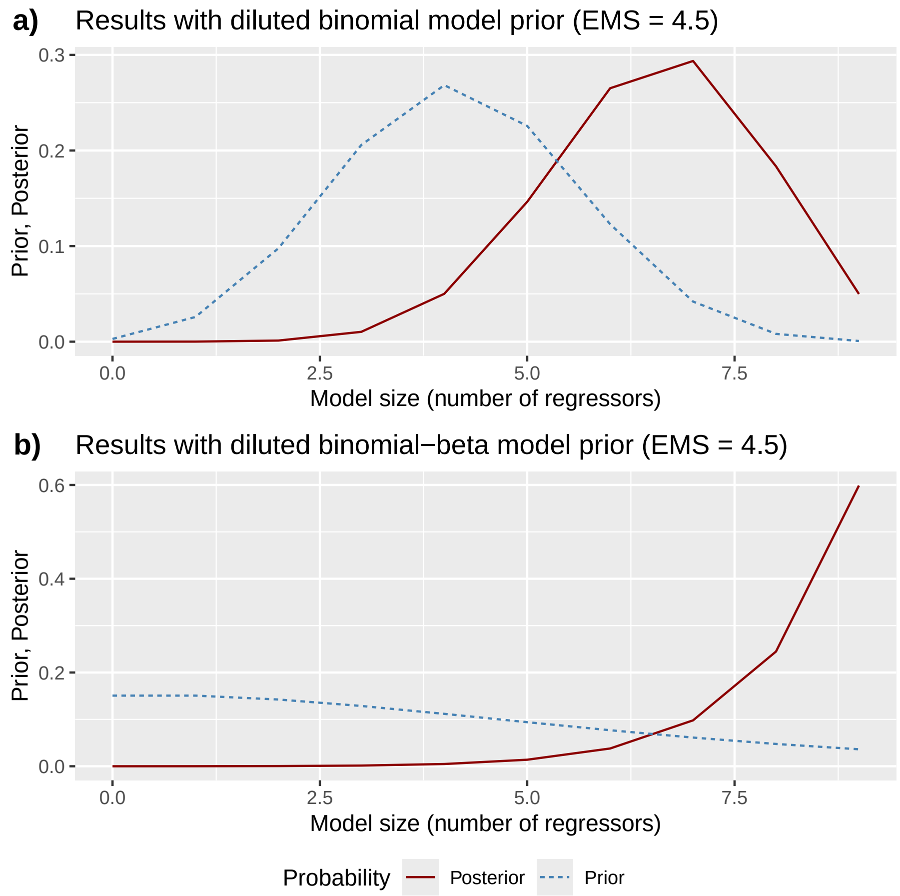
One of the main issues associated with identifying robust regressors is multicollinearity. Some regressors may approximate the same underlying factor influencing the dependent variable. Multicollinearity may result from the absence of observable variables associated with a specific theory or from a theory failing to provide a unique candidate for a regressor. Moreover, some regressors may share a common determinant. Although Moral-Benito (2013, 2016) addressed this issue to some extent, researchers have another option to mitigate multicollinearity: the dilution prior proposed by George (2010) which was described in detail in subsection 2.4.

To apply the dilution prior, the user must set `dilution = 1` in the `bma` function. The user can also manipulate the dilution parameter $\omega$. The default option is `dil.Par = 0.5`, as recommended by George (2010).

```
> bma_results_dil <- bma(bma_prep_objects_full, df = data_prepared,
+                        round = 3, dilution = 1)
```

The effect of implementing the dilution prior is well depicted by the distribution of prior probability mass over the model sizes.

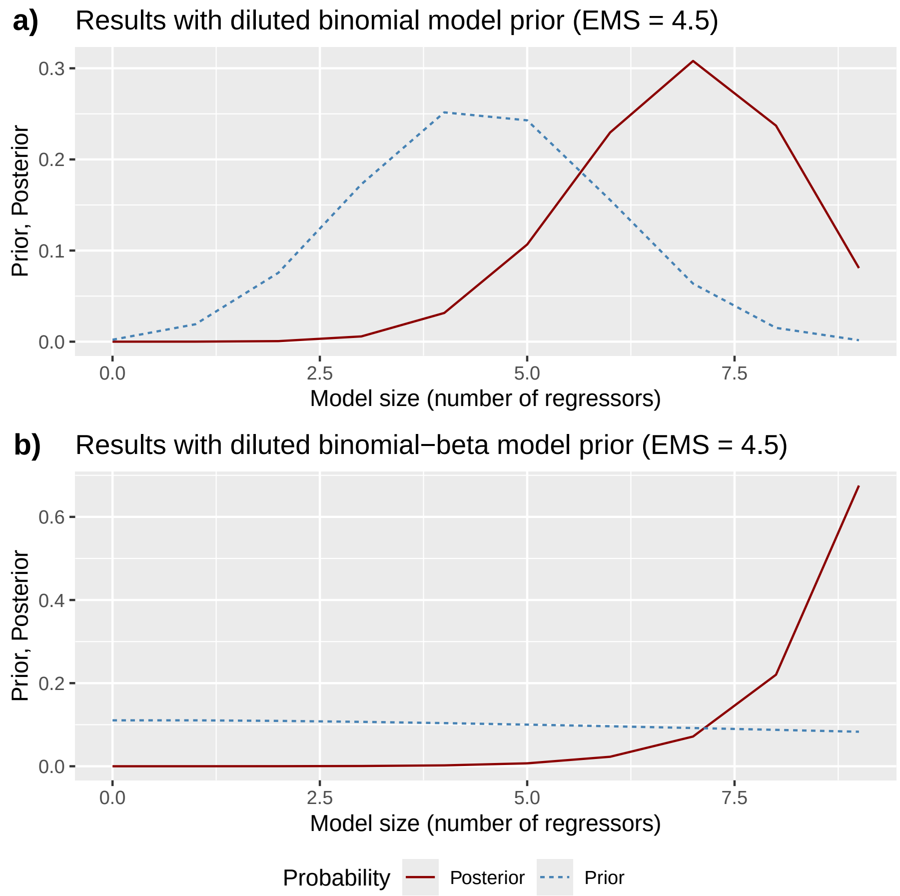
```
> size_graphs_dil <- model_sizes(bma_results_dil)
```




The change in the prior distribution is more visible for the binomial-beta model prior. In panel b, the prior probability mass has decreased for larger models and increased for smaller models. However, this change is not uniform, as models characterized by the highest degree of multicollinearity are subject to the greatest penalty in terms of prior probability mass.

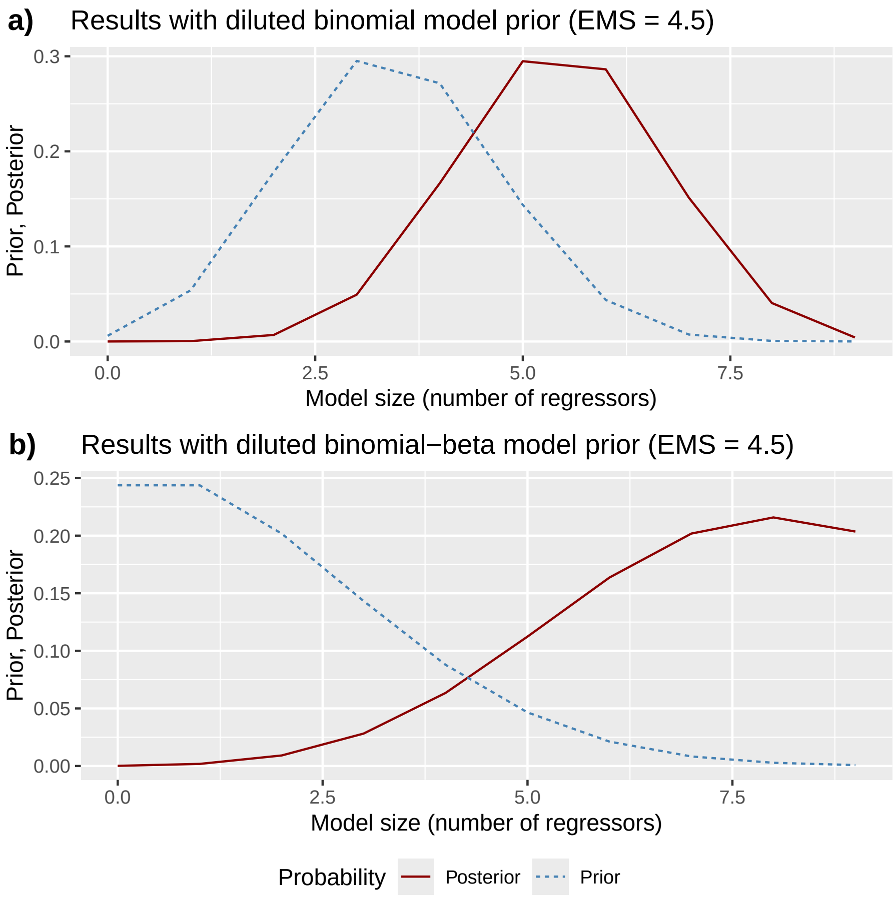
Before moving to the BMA statistics, it is instructive to examine the change in the `dil.Par` parameter.

```
> bma_results_dil01 <- bma(bma_prep_objects_full, df = data_prepared,
+                          round = 3, dilution = 1, dil.Par = 0.1)
> size_graphs_dil01 <- model_sizes(bma_results_dil01)
```




As we can see, decreasing the value of $\omega$ diminishes the impact of dilution on the model prior. Conversely, raising the `dil.Par` parameter increases the degree of dilution.

```
> bma_results_dil2 <- bma(bma_prep_objects_full, df = data_prepared,
+                         round = 3, dilution = 1, dil.Par = 2)
> size_graphs_dil2 <- model_sizes(bma_results_dil2)
```




An especially strong impact can be seen for the binomial-beta prior.

However, even after giving such priority to the penalty for multicollinearity, the main BMA statistics remain stable.

```
> bma_results_dil2[[2]]

          PIP     PM   PSD  PSDR  PMcon PSDcon PSDRcon    %(+)
gdp_lag    NA  0.924 0.076 0.107  0.924  0.076   0.107 100.000
ish     0.735  0.056 0.044 0.060  0.076  0.033   0.058 100.000
sed     0.641  0.029 0.054 0.067  0.046  0.062   0.078  70.312
pgrw    0.687  0.019 0.030 0.052  0.028  0.033   0.060  99.609
pop     0.993  0.121 0.064 0.080  0.122  0.063   0.080 100.000
ipr     0.773 -0.039 0.032 0.044 -0.050  0.027   0.044   0.000
opem    0.824  0.037 0.029 0.031  0.045  0.026   0.029 100.000
gsh     0.840 -0.014 0.042 0.094 -0.017  0.045   0.102  28.906
lnlex   0.767  0.086 0.075 0.097  0.112  0.066   0.097 100.000
polity  0.613 -0.050 0.046 0.052 -0.081  0.030   0.043   0.000
```

Hence, we see that Moral-Benito (2016)'s claim about the fragility of growth regressors withstands the test of various manipulations in the model prior.

# 7 Concluding remarks

This manuscript introduces the `bdsm` package, which enables Bayesian model averaging for dynamic panels with weakly exogenous regressors — a methodology developed by Moral-Benito (2012, 2013, 2016). This package allows researchers to simultaneously address model uncertainty and reverse causality and is the only `R` package offering these capabilities. It provides flexible options for specifying model priors, including dilution prior that accounts for multicollinearity. The package also includes graphical tools for visualizing prior and posterior model probabilities across model space and model sizes, as well as functions for plotting histograms and kernel densities of the estimated coefficients. Additionally, it allows researchers to compute jointness measures introduced by Doppelhofer & Weeks (2009); Ley & Steel (2007); Hofmarcher et al. (2018) to assess whether pairs of regressors act as substitutes or complements. Users can also perform Bayesian model selection to examine in detail the most probable models based on posterior model probability.

The manuscript outlines the methodological approach, while the detailed explanation can be found in Moral-Benito (2012, 2013, 2016). Users unfamiliar with this approach can easily learn to apply it through the hands-on tutorial provided in the manuscript. The package's functionalities are illustrated using the original dataset from Moral-Benito (2016) in the context of analyzing the determinants of economic growth. The results of the examination illustrate that fragility of growth determinants is a persistent feature of the data, confirming Moral-Benito (2016) claims. The various empirical exercises underscore two important aspects of any BMA analysis. First, the results should always be validated through extensive changes in prior specifications. Second, the robustness of the regressors must be evaluated using both posterior inclusion probabilities and the ratios of the posterior mean to the posterior standard deviation, as these measures can often lead to differing conclusions.

# References


Aller, C., Ductor, L., & Grechyna, D. (2021). Robust determinants of co2 emissions. *Energy Economics*, 96(105154).

Amini, S. M. & Parmeter, C. F. (2011). Bayesian model averaging in r. *Journal of Economic and Social Measurement*, 36(4), 253–287.

Arin, K. P., Braunfels, E., & Doppelhofer, G. (2019). Revisiting the growth effects of fiscal policy: A Bayesian model averaging approach. *Journal of Macroeconomics*, 62(103158), 1–16.

Baran, S. & Möller, A. (2015). Joint probabilistic forecasting of wind speed and temperature using bayesian model averaging. *Environmetrics*, 26(2), 120–132.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, 106(2), 407–443.

Beck, K. (2017). Bayesian model averaging and jointness measures: theoretical framework and application to the gravity model of trade. *Statistics in Transition. New Series*, 18(3), 393–412.

Beck, K. (2022). Macroeconomic policy coordination and the European business cycle: Accounting for model uncertainty and reverse causality. *Bulletin of Economic Research*, 74(4), 1095–1114.

Błażejowski, M. & Kwiatkowski, J. (2015). Bayesian Model Averaging and Jointness Measures for gretl. *Journal of Statistical Software*, 68(5), 1–24.

Chen, H., Mirestean, A., & Tsangarides, C. G. (2018). Limited information bayesian model averaging for dynamic panels with application to a trade gravity model. *Econometric Reviews*, 37(7), 777–805.

Clyde, M. A., Ghosh, J., & Littman, M. L. (2011). Bayesian adaptive sampling for variable selection and model averaging. *Journal of Computational and Graphical Statistics*, 20(1), 80–101.

D'Andrea, S. (2022). Are there any robust determinants of growth in europe? a bayesian model averaging approach. *International Economics*, 171, 143–173.

Doppelhofer, G. & Weeks, M. (2009). Jointness of growth determinants. *Journal of Applied Econometrics*, 24(2), 209–244.

Ductor, L. & Leiva-Leon, D. (2016). Dynamics of Global Business Cycles Interdependence. *Journal of International Economics*, 102, 1109–1127.

Eicher, T. S., Papageorgiou, C., & Raftery, A. E. (2011). Default priors and predictive performance in Bayesian model averaging, with application to growth determinants. *Journal of Applied Econometrics*, 26(1), 30–55.

Eicher, T. S., Papageorgiou, C., & Roehn, O. (2007). Unraveling the fortunes of the fortunate: An Iterative Bayesian Model Averaging (IBMA) approach. *Journal of Macroeconomics*, 29(3), 494–514.

Feldkircher, M. & Zeugner, S. (2015). Bayesian Model Averaging Employing Fixed and Flexible Priors: The BMS Package for R. 68(4), 1–37.

Fernández, C., Ley, E., & Steel, M. F. (2001a). Benchmark priors for Bayesian model averaging. *Journal of Econometrics*, 100(2), 381–427.

Fernández, C., Ley, E., & Steel, M. F. (2001b). Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics*, 16(5), 563–576.

Figini, S. & Giudici, P. (2017). Credit risk assessment with bayesian model averaging. *Communications in Statistics - Theory and Methods*, 46(19), 9507–9517.

Fragoso, T. M., Bertoli, W., & Louzada, F. (2018). Bayesian model averaging: A systematic review and conceptual classification. *International Statistical Review*, 86(1), 1–28.

George, E. I. (2010). Dilution priors: Compensating for model space redundancy. In J. O. Berger, T. Cai, & I. M. Johnstone (Eds.), *Borrowing Strength: Theory Powering Applications – A Festschrift for Lawrence D. Brown* (pp. 158–165). Beachwood, OH: Institute of Mathematical Statistics.

Granger, C. W. & Uhlig, H. F. (1990). Reasonable extreme-bounds analysis. *Journal of Econometrics*, 44(1), 159–170.

Guliyev, H. (2024). Determinants of ecological footprint in european countries: Fresh insight from bayesian model averaging for panel data analysis. *Science of The Total Environment*, 912, 169455.

Hlavac, M. (2016). Extremebounds: Extreme bounds analysis in r. *Journal of Statistical Software*, 72(9), 1–22.

Hofmarcher, P., Crespo Cuaresma, J., Grün, B., Humer, S., & Moser, M. (2018). Bivariate joint measure in Bayesian Model Averaging: Solving the conundrum. *Journal of Macroeconomics*, 57, 150–165.

Horvath, R., Horvatova, E., & Siranova, M. (2024). The determinants of financial development: Evidence from bayesian model averaging. *Economic Systems*, (101274).

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 186(1007), 453–461.

Kass, R. E. & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795.

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

Leamer, E. E. (1983). Let's Take the Con Out of Econometrics. *The American Economic Review*, 73(1), 31–43.

Leamer, E. E. (1985). Sensitivity Analyses Would Help. *American Economic Review*, 75(3), 308–313.

Leamer, E. E. & Leonard, H. (1981). Reporting the fragility of regression estimates. *The Review of Economics and Statistics*, 65(2), 306–317.

Lenkoski, A., Eicher, T. S., & Raftery, A. E. (2014). Two-stage bayesian model averaging in endogenous variable models. *Econometric Reviews*, 33(1-4), 122–151.

Levine, R. & Renelt, D. (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *The American Economic Review*, 82(4), 942–963.

Ley, E. & Steel, M. F. (2007). Jointness in Bayesian variable selection with applications to growth regression. *Journal of Macroeconomics*, 29(3), 476–493.

Ley, E. & Steel, M. F. (2009). On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics*, 24(4), 651–674.

Ley, E. & Steel, M. F. (2012). Mixtures of g-priors for Bayesian model averaging with economic applications. *Journal of Econometrics*, 171(2), 251–266.

León-González, R. & Montolio, D. (2015). Endogeneity and panel data in growth regressions: A Bayesian model averaging approach. *Journal of Macroeconomics*, 24, 23–39.

Liu, C. & Maheu, J. M. (2009). Forecasting realized volatility: a bayesian model-averaging approach. *Journal of Applied Econometrics*, 24(5), 709–733.

Masanjala, W. H. & Papageorgiou, C. (2008). Rough and lonely road to prosperity: a reexamination of the growth in Africa using Bayesian model averaging. *Journal of Applied Econometrics*, 23(5), 671–682.

Mirestean, A. & Tsangarides, C. G. (2016). Growth Determinants Revisited Using Limited-Information Bayesian Model Averaging. *Journal of Applied Econometrics*, 31(1), 106–132.

Moral-Benito, E. (2012). Determinants of Economic Growth: A Bayesian Panel Data Approach. *The Review of Economics and Statistics*, 92(4), 566–579.

Moral-Benito, E. (2013). Likelihood-Based Estimation of Dynamic Panels with Predetermined Regressors. *Journal of Business and Economic Statistics*, 31(4), 451–472.

Moral-Benito, E. (2015). Model Averaging in Economics: An Overview. *Journal of Economic Surveys*, 29(1), 46–75.

Moral-Benito, E. (2016). Growth Empirics in Panel Data Under Model Uncertainty and Weak Exogeneity. *Journal of Applied Econometrics*, 31(3), 584–602.

Moral-Benito, E., Allison, P., & Williams, R. (2019). Dynamic panel data modelling using maximum likelihood: An alternative to Arellano-Bond. *Applied Economics*, 51(20), 2221–2232.

Moser, M. & Hofmarcher, P. (2014). Model priors revisited: Interaction terms in BMA growth applications. *Journal of Applied Econometrics*, 29(2), 344–347.

Payne, R. D., Ray, P., & Thomann, M. A. (2024). Bayesian model averaging of longitudinal dose-response models. *Journal of Biopharmaceutical Statistics*, 34(3), 349–365.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163.

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92(437), 179–191.

Raftery, A. E., Painter, I. S., & Volinsky, C. T. (2005). BMA: An R package for bayesian model averaging. *R News*, 5(2), 2–8.

Sala-I-Martin, X. (1997). I Just Ran Two Million Regressions. *The American Economic Review*, 27(2), 178–183.

Sala-I-Martin, X., Doppelhofer, G., & Miller, R. I. (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *The American Economic Review*, 94, 813–835.

Sloughter, J. M., Gneiting, T., & Raftery, A. E. (2013). Probabilistic wind vector forecasting using ensembles and bayesian model averaging. *Monthly Weather Review*, 141(6), 2107–2119.

Steel, M. F. (2020). Model Averaging and Its Use in Economics. *Journal of Economic Literature*, 58(3), 644–719.