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

Krzysztof Beck
Lazarski University

Marcin Dubel
Lazarski University

Mateusz Wyszyński
Lazarski University

March 2, 2025

Abstract

Place for abstract

Contents

1 Introduction

Raftery et al. (2005)
Feldkircher and Zeugner (2015)
Błażejowski and Kwiatkowski (2015)

Leamer (1978)
Leamer (1983)
Leamer (1985)

León-González and Montolio (2015)
Mirestean and Tsangarides (2016)
Chen et al. (2018)

Moral-Benito (2015)
Steel (2020)

2 Model setup and Bayesian model averaging

2.1 Model setup

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

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

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, β is a parameter vector, η_i is an entity specific fixed effect, ζ_t is a period-specific shock and v_{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

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

where $y_t^{t-1} = (y_{ij,0}, ..., y_{ij,t-1})'$ and $x_t = (x_{ij,0}, ..., x_{ij,t})'$. 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 a reduced-form equation 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 $h < t$, γ_{th} is a $k \times 1$ vector $(y_{th}^1, \ldots, y_{th}^k)'$ $h = 0, \ldots, T-1$; Λ_{th} is a $k \times k$ matrix of parameters, and ϑ_{it} is a $k \times 1$ of prediction errors. The mean vector and the covariance matrix of the joint distribution of the initial observations and the individual effects η_i are unrestricted:

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

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

where c_0 is a scalar, and c_1 and γ_{10} are $k \times 1$ vectors.[1]

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 θ denotes parameters to be estimated, $R_i = (y_{io}, x_{i1}', y_{i1}, \ldots, x_{iT}', y_{iT})'$ are vectors of observed variables, and $\Sigma = diag[\sigma_\eta^2, \sigma_{v_0}^2, \Sigma_{\vartheta_1}, \sigma_{v_1}^2, ..., \Sigma_{\vartheta_T}, \sigma_{v_T}^2]$ is the block-diagonal variance-

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

covariance matrix. Matrix B is given by:

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

and matrix D is given by:

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

The model setup in equations (1) and (3-5) requires that in addition to the parameters of interest α and β, the parameters γ_{ij} and Λ_{km} need to be estimated. To make the optimisation 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:

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

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 α, β, ϕ_0, ϕ_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:

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

where U_1 is a matrix of errors connected only to dependent variables, Σ_{11} is a part of the Σ matrix:

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

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 $F_{12} = -\Sigma_{11}^{-1}\Sigma_{12}$.

2.2 Bayesian model averaging

Given the likelihood function in (17), henceforth denoted as $L(\text{data}|\theta_i, M_j)$ for a specific model i, it is possible to utilize Bayesian model averaging[2] (BMA). To achieve that, we first estimate all possible variants of equation:

$$Y = X\beta + \epsilon \tag{19}$$

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

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

where $P(M_j)$ 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:

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

Finally, using the result that

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

we can calculate posterior model probability as

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

[2] For an introduction to BMA see Raftery (1995); Raftery et al. (1997); Fernández et al. (2001a,b); Doppelhofer and Weeks (2009); Beck (2017).

2.3 BMA statistics

With PMPs, it is possible to calculate BMA statistics. The probability that a given variable should be included in a model after seeing the data, i.e. the posterior inclusion probability (PIP) is given by:

$$P(x_k|\text{data}) = \sum_{j=1}^{2^K} 1(\pi_k = 1|\text{data}, M_j) \times P(M_j|\text{data}) \tag{28}$$

where x_k is regressor k ($k = 1, \ldots, K$), π_j is $(K \times 1)$ binary vector, where 1 indicates presence of a regressor in a model, and $1(\pi_j|\text{data}, M_j)$ is an indicator function that takes the value 1 when regressor x_k is present in model j and 0 otherwise. Posterior mean (PM) is calculated as:

$$E(\beta_k|\text{data}) = \sum_{j=1}^{2^K} \widehat{\beta}_{k,j} \times P(M_j|\text{data}) \tag{29}$$

where $\widehat{\beta}_{k,j}$ is the value of the coefficient β_k in model j. The posterior standard deviation (PSD) is equal to:

$$SD(\beta_k|\text{data}) = \sqrt{\sum_{j=1}^{2^K} V(\beta_{k,j}|\text{data}, M_j) \times P(M_j|\text{data}) + \sum_{j=1}^{2^K} [\widehat{\beta}_{k,j} - E(\beta_k|\text{data})]^2 \times P(M_j|\text{data})} \tag{30}$$

where $V(\beta_{k,j}|data, M_j)$ denotes the conditional variance of the coefficient β_k in model M_j.

Alternatively, one might be interested in the values of the coefficients and variances on the condition of inclusion of a given regressor in a model. The conditional posterior mean (PM_con) is given by:

$$E(\beta_k|\pi_j = 1, \text{data}) = \frac{E(\beta_k|\text{data})}{P(x_k|\text{data})}. \tag{31}$$

The conditional standard deviation (PSD_con) is:

$$SD(\beta_k|\pi_j = 1, \text{data}) = \sqrt{\frac{SD(\beta_k|\text{data})^2 + E(\beta_k|\text{data})^2}{P(x_k|\text{data})} - E(\beta_k|\pi_j = 1, \text{data})^2}. \tag{32}$$

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 and 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[3]. The package offers two main options. The first is binomial model prior (Sala-I-Martin et al., 2004):

$$P(M_j) = \left(\frac{EMS}{K}\right)^{k_j} \left(1 - \frac{EMS}{K}\right)^{K-k_j} \tag{33}$$

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

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

$$P(M_j) \propto \Gamma(1 + k_j)\Gamma(\frac{K - EMS}{EMS} + K - k_j). \tag{34}$$

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

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

$$P_D(M_j) \propto P(M_j)|COR_j|^{\omega} \tag{35}$$

where $P_D(M_j)$ is the diluted model prior, $|COR_j|$ is the determinant of the correlation matrix of regressors in model j, and ω 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[4]. Assuming two different covariates a and b, let $P(a \cap b)$ be the posterior probability of the inclusion of both variables, $P(\overline{a} \cap \overline{b})$ the posterior probability of the exclusion of both variables, $P(\overline{a} \cap b)$ and $P(a \cap \overline{b})$ denote the posterior probability of including each variable separately. The first measure of jointness is simply $P(a \cap b)$. However, this measure ignores much of the information about the relationships between the regressors. Doppelhofer and Weeks (2009) measure if defined as:

$$J_{DW} = \log \left[\frac{P(a \cap b) * P(\overline{a} \cap \overline{b})}{P(\overline{a} \cap b) * P(a \cap \overline{b})} \right]. \tag{36}$$

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 and Steel (2007) is given by:

$$J_{LS} = \frac{P(a \cap b)}{P(\overline{a} \cap b) + P(a \cap \overline{b})}. \tag{37}$$

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:

$$J_{HCGHM} = \frac{(P(a \cap b) + \rho) * P(\overline{a} \cap \overline{b}) + \rho) - (P(\overline{a} \cap b) + \rho) * P(a \cap \overline{b}) + \rho)}{(P(a \cap b) + \rho) * P(\overline{a} \cap \overline{b}) + \rho) + (P(\overline{a} \cap b) + \rho) * P(a \cap \overline{b}) + \rho) + \rho}. \tag{38}$$

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.

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", repos = "https://cloud.r-project.org")
```

```
> library(bdsm)
```

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

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) and time. 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 transformations using the `data_prep` function. The user might want to prepare the data for a fixed effects model. For time fixed effects, the user can perform cross-sectional demeaning (`time_effects = TRUE`), while for cross-section effects, the user can perform time demeaning (`entity_effects = TRUE`). Moreover, the user can scale the data by the standard deviation within cross-sections (`time_scale = TRUE`) and within time periods (`entity_scale = TRUE`). The user can also perform regular standardization by subtracting the mean from each column (`standardize = TRUE`) and dividing it by the standard deviation (`scale = TRUE`). Standardization is preferred because it improves the computational efficiency of the likelihood optimization. Finally, the user can specify the order in which demeaning and/or scaling should be applied using the `order` parameter. This parameter takes a vector with three elements, where the order of the elements determines the sequence of operations. Here, `S`, `T`, `E`, and `0` denote standardization, preparation for time effects, preparation for entity effects, and skipping an operation, respectively. For example, to apply the entity effect first and then standardization, use:

```
> data_prepared <- data_prep(df = economic_growth, timestamp_col = year,
+                            entity_col = country, standardize = TRUE,
+                            scale = TRUE, entity_effects = TRUE,
+                            order = c("T", "S", "0"))
```

Moral-Benito (2016) first standardized the regressors, leaving the dependent variable unchanged, and then introduced time fixed effects. To achieve this effect, use:

```
> data_prepared <- economic_growth |>
+     data_prep(timestamp_col = year, entity_col = gdp,
+               order = c("S","0","0")) |>
+     data_prep(timestamp_col = year, entity_col = country,
+               standardize = FALSE, scale = FALSE,
+               time_effects = TRUE, time_scale = FALSE,
+               order = c("T","0","0"))
```

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

```
Example output
```

4 Usage

4.1 Basic Workflow

```
R> library(bdsm)
R> data(mydata)
R> result <- bma(y ~ ., data = mydata)
R> summary(result)
```

```
# Output here
```

5 Technical Details

- Implements g-prior for Bayesian Linear Regression.
- Supports model comparison and posterior analysis.

6 References

References

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.

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

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

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

Eicher, T. S., Papageorgiou, C., and 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.

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

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

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

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

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

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.

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

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.

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

Ley, E. and 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.

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

Masanjala, W. H. and 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. and 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., and Williams, R. (2019). Dynamic panel data modelling using maximum likelihood: An alternative to Arellano-Bond. *Applied Economics*, 51(20):2221–2232.

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

Raftery, A. E., Madigan, D., and 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., and Volinsky, C. T. (2005). BMA: An R package for bayesian model averaging. *R News*, 5(2):2–8.

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

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