# Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies


MASTER THESIS

# Spread Determinants and Model Uncertainty: A Bayesian Model Averaging Analysis


Author: **Bc. Vojtěch Seman**

Supervisor: **PhDr. Marek Rusnák**


Academic Year: **2010/2011**

## Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 31, 2011

Signature

## Acknowledgments

The author is thankful especially to the thesis supervisor for his fruitful comments, suggestions and help. The author is also grateful to his family for their support.

## Abstract

The spread between interest rate and sovereign bond rate is commonly used indicator for country's probability to default. Existing literature proposes many different potential spread determinants but fails to agree on which of them are important. As a result, there is a considerable uncertainty about the correct model explaining the spread. We address this uncertainty by employing Bayesian Model Averaging method (BMA). The BMA technique attempts to consider all the possible combinations of variables and averages them using a model fit measure as weights. For this empirical exercise, we consider 20 different explanatory variables for a panel of 47 countries for the 1980-2010 period. Most of the previously suggested determinants were attributed high inclusion probabilities. Only the "foreign exchange reserves growth" and the "exports growth" scored low by their inclusion probabilities. We also find a role of variables previously not included in the literature's spread determinants — "openness" and "unemployment" which rank high by the inclusion probability. These results are robust to a wide range of both parameter and model priors.

| | |
|---|---|
| **JEL Classification** | C6, C8, C11, C51, E43 |
| **Keywords** | Sovereign Spread Determinants, Model Uncertainty, Bayesian Model Averaging |

| | |
|---|---|
| **Author's e-mail** | `semanv()gmail()com` |
| **Supervisor's e-mail** | `marek.rusnak()cerge-ei()cz` |

# Abstrakt

Rozdíl mezi úrokovými sazbami a sazbou vládních dluhopisů patří mezi běžně používaný indikátor pravděpodobnosti státního bankrotu — tzv. spread. Stávající literatura jej vyjadřuje pomocí mnoha determinant, ale neshoduje se v tom, které z těchto determinant jsou považovány za důležité. Z tohoto důvodu panuje značná nejistota o tom, který z možných modelů je tím pravým. Tento problém řešíme pomocí metody "Bayesian Model Averaging" (BMA), která tomuto problému předchází tím, že vyhodnocuje míru pravděpodobnosti jednotlivých modelů, a jako váhy je používá pro průměrování. Pro toto empirické cvičení používáme panel 47 zemí s dvaceti vysvětlujícími proměnnými pro období 1980-2010. Pro většinu determinant z literatury byly odhadnuty vysoké míry pravděpodobnosti pro jejich začlenění do modelu. Pro "růst rezerv cizí měny" a "nárůst exportu" byly tyto míry zjištěny nízké. Naopak proměnným "otevřenost ekonomiky" a "nezaměstnanost" byly přiděleny vysoké míry pravděpodobnosti, ač mezi literaturou doporučované determinanty spreadu nepatří. Tyto výsledky jsou robustní vůči škále alternativních předpokladů o priorních distribucích (BMA priors) a to jak na modelový, tak i na parametrický prostor.



| | |
|---|---|
| **Klasifikace JEL** | C6, C8, C11, C51, E43 |
| **Klíčová slova** | Determinanty spreadu vládních dluhopisů, Modelová nejistota, Bayesovské modelové průměrování |



| | |
|---|---|
| **E-mail autora** | semanv()gmail()com |
| **E-mail vedoucího práce** | marek.rusnak()cerge-ei()cz |

# Contents

# List of Tables

# List of Figures

# Acronyms

**BMA** Bayesian Model Averaging

**CEE** Central and Eastern Europe

**EIU** Economist Intelligence Unit

**EMBIG** Emerging Market Bond Index Global

**FDI** Foreign Direct Investment

**GDP** Gross Domestic Product

**IFS** International Financial Statistics

**iid** independent and identically distributed

**IMF** International Monetary Fund

**MCMC** Monte Carlo Markov Chain

**ML** Marginal Likelihood

**OECD** Organisation for Economic Co-operation and Development

**OLS** Ordinary Least Squares

**PIP** Posterior Inclusion Probability

**PMP** Posterior Model Probability

# Chapter 1

# Introduction

A spread between interest rate and sovereign bond rate is commonly used as a proxy for an economy's creditworthiness (Rowland & Torres 2004), as a country's probability to default (Baldacci 2007; della Paolera & Grandes 2007) and is also linked to sovereign debt pricing (Hilscher & Nosbusch 2010), sovereign ratings (Kamin & von Kleist 1999; McGuire & Schrijvers 2003; Uribe & Yue 2006) and market sentiments (Eichengreen & Mody 1998).

For the selection of spread determinants, we combine the spread literature (Rowland & Torres 2004; Cantor & Packer 1996) with related Early Warning Systems literature (Alessi & Detken 2009; Frankel & Saravelos 2010; Rose & Spiegel 2011). The variables from both "camps" overlap and the existing literature proposes many different potential determinants.

As the most important variables, some studies identify the *exchange rate changes* (Frankel & Rose 1996) to show external country wealth which serves as a central bank tool to eliminate the currency market volatility. Others propose *reserves movement* to account the importance of the liquidity of an economy (Rose & Spiegel 2011). Others show as the most important variable the *interest rate changes* and both of the formerly stated variables as a central bank instrument to manage a defense against speculative attacks (Hawkins & Klau 2000; Eichengreen October 1995). Other studies find the most significant determinants in the *indebtedness* and *current account balance* (Edwards 1984) and other in the *inflation* (Min 1998).

Overall, the literature fails to agree on which of the spread determinants are important (Frankel & Saravelos 2010; Rowland 2004; Rowland & Torres 2004; Baldacci *et al.* 2008). As a result, there is a considerable uncertainty about the true model explaining the spread.

We address this uncertainty by employing the Bayesian Model Averaging (BMA) method, similarly with Moral-Benito (2010a) and Crespo Cuaresma & Slacik (2009). The BMA technique attempts to consider all the possible combinations of variables and averages them using a model fit measure as weights.

Despite the BMA being a "robust" method, a special attention needs to be paid to the transformation of determinants (Doppelhoffer & Weeks 2008). Finally, for this empirical exercise, we consider 20 different explanatory variables for a panel of 47 countries for the period between 1980 and 2010.

The problematics of the true spread determinants is also handled in Obstfeld *et al.* (2009) and Rose & Spiegel (2009). There are also existing overviews providing occurrence count for separate spread determinants across studies (Hawkins & Klau 2000; Kaminsky *et al.* 1998; Abiad 2003). By aggregating these three overviews together, Frankel & Saravelos (2010) create a rank of variables based on the occurrence count. In this BMA study, we challenge the results with the findings of Rowland & Torres (2004), Cantor & Packer (1996) and Frankel & Saravelos (2010).

We employ a sensitivity analysis of the BMA priors (Fernandez *et al.* 2001; Feldkircher & Zeugner 2009; Eicher *et al.* 2011) to check the robustness of results.

The objective of this thesis is twofold. First, we want to select and rank the variables according to their likeliness to be included into the true model describing the sovereign spread, controlling for the model uncertainty. Second, to be able to do the step above, we present the BMA method.

This thesis is structured as follows. Chapter 2 describes the dataset we use. Chapter 3 or "from the Bayes theorem to the probability of a variable's inclusion", presents the BMA methodology and its items step by step. Chapter 4 provides the information necessary to build a fixed effects model using the BMA method. Results are shown in the Chapter 5. The robustness of the results is tested by various options the BMA method offers in the Chapter 6. Chapter 7 concludes.

# Chapter 2

# Dataset

Our dataset is inspired by several studies concerning both the determinants of the sovereign spreads, the country crisis vulnerability measure and the Early Warning Systems' indicators. Namely, we draw the inspiration from Rowland (2004), Frankel & Saravelos (2010), Baldacci *et al.* (2008), Alessi & Detken (2009), and Rose & Spiegel (2009).

We use a panel of data for the empirical analysis of the spread between sovereign bond yield and a 3 month money market rate. Our panel consists of 21 financial and macroeconomic variables described in the Table 2.2. The panel comprises 47 countries that are listed in the Table 2.1[1]. The panel consists of yearly data for latest 31 years, ranging from 1980 to 2010. As some of the countries included in the panel either did not existed for the whole observed period or did not report the data, the panel is unbalanced.

The source of data is the Economist Intelligence Unit (EIU) online database. However, their original source are international financial institutions like the International Monetary Fund (IMF)'s International Financial Statistics (IFS) and the Worldbank or national statistical units and national central banks.

All variables were transformed appropriately to ensure the data stationarity. The variables marked *dl* in their names were transformed to be used as yearly growth rates using a difference of natural logarithms. Some variables were transformed as a share on the Gross Domestic Product (GDP) which can be noted from the description column in the Table 2.2.

As the BMA method requires full rank of variables' vectors for a specific

[1]The only key to select countries is the data availability, the focus is worldwide. Most of the Organisation for Economic Co-operation and Development (OECD) countries are covered, while the data are not available only for these OECD countries Chile, Iceland, Israel, Luxembourg and Turkey in the Economist Intelligence Unit (EIU) database.

| | | |
|---|---|---|
| Australia | Hong Kong | Peru |
| Austria | Hungary | Philippines |
| Belgium | India | Poland |
| Brazil | Indonesia | Portugal |
| Bulgaria | Ireland | Romania |
| Canada | Italy | Russian Federation |
| China | Japan | Singapore |
| Colombia | Kazakhstan | Slovakia |
| Czech Republic | Korea, Rep. Of | Slovenia |
| Denmark | Latvia | Spain |
| Ecuador | Lithuania | Sweden |
| Estonia | Malaysia | Switzerland |
| Finland | Mexico | Taiwan |
| France | Netherlands | United Kingdom |
| Germany | New Zealand | United States |
| Greece | Norway | |

Table 2.1: List of countries

country and year across all variables. If one of the variable fields is missing, the whole year for such country is not taken into account. That is why the data inspection was necessary. Before downloading the final data, we checked the data availability across all variables accessible at the EIU statistics, for every country available and across latest 40 years. Finally, after the data availability inspection, the full data matrix was observed for several relevant variables, for the most of the developed countries of which all of the OECD countries and for the time period between 2001 and 2010 which was taken into account. To achieve larger variability in the data we extended the observed period starting by 1980, expecting some missing data in the dataset[2].

[2]We found a minor inconsistency in the data set using the variables' summary statistics. Further inspection led to replace the values provided for the Russian Federation in the beginning of its time series for 1990-1993. These values for 4 observations in TDRA, 1 observation in BEXP and 1 in PSBR were replaced by the average value of remaining years. In another case, the values of CCPI and LRAT in Peru during hyperinflation between 1989 and 1990 were topped by 200 %. The replaced observations are considered as outliers whose presence is undesirable for the results (increasing the variance and reducing the explaining power).

| Variable Name | Description | Unit | Transformation |
|---|---|---|---|
| spread | 10yr sov debt yield - 3M IBOR | percent | 0 |
| BEXP | budget expenditure / GDP | percent | 0 |
| DGDP | real GDP growth | percent | 0 |
| INVR | foreign direct investment / GDP | percent | 0 |
| LRAT | lending interest rate | percent | 0 |
| NBTT | terms of trade | percent | 0 |
| PSBR | budget balance / GDP | percent | 0 |
| PUDP | public debt / GDP | percent | 0 |
| TDRA | trade balance / GDP | percent | 0 |
| UNEM | unemployment | percent | 0 |
| dlBEXL | budget expenditure growth | percent | 1 |
| dlCCPI | inflation | percent | 1 |
| dlCEXP | exports growth | percent | 1 |
| dlFRES | foreign exchange reserves excl.gold) growth | percent | 1 |
| dlLCHD | labor cost growth | percent | 1 |
| dlMIPD | abroad debt service growth | percent | 1 |
| dlRIND | real industry growth | percent | 1 |
| dlXRPD | exchange rate (domestic/USD) change | percent | 1 |
| fresgdp | foreign exchange reserves (excl.gold) / GDP | percent | 0 |
| open | openness ratio ($\frac{X+M}{2GDP}$) | ratio | 0 |
| xrdev | exchange rate deviation from trend | percent | 2 |

*Note:* transformation 0 stands for none, 1 for difference of logarithms and 2 for deviation from trend (1980-2010). $IBOR$ stands for interbank offered rate.

Table 2.2: Description of variables

| Variable | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| spread | 1.05 | 3.36 | -35.14 | 39.41 |
| BEXP | 36.52 | 13.74 | 10.06 | 79.78 |
| DGDP | 3.21 | 4.19 | -21.26 | 24.62 |
| INVR | 3.02 | 4.30 | -15.00 | 36.62 |
| LRAT | 14.62 | 26.50 | 1.38 | 644.50 |
| NBTT | 102.86 | 19.05 | 36.03 | 233.70 |
| PSBR | -2.13 | 4.57 | -40.87 | 19.26 |
| PUDP | 50.43 | 30.19 | 1.30 | 197.53 |
| TDRA | 0.20 | 8.26 | -64.25 | 30.25 |
| UNEM | 7.66 | 4.03 | 0.19 | 24.13 |
| dlBEXL | 0.10 | 0.16 | -0.59 | 1.90 |
| dlCCPI | 0.10 | 0.27 | -0.05 | 3.08 |
| dlCEXP | 0.06 | 0.08 | -0.51 | 0.41 |
| dlFRES | 0.10 | 0.38 | -3.25 | 4.81 |
| dlLCHD | 0.06 | 0.13 | -0.92 | 1.00 |
| dlMIPD | 0.11 | 0.25 | -2.33 | 2.30 |
| dlRIND | 0.03 | 0.06 | -0.41 | 0.63 |
| dlXRPD | 0.06 | 0.31 | -0.35 | 4.25 |
| fresgdp | 0.12 | 0.18 | 0.00 | 1.31 |
| open | 0.38 | 0.32 | 0.05 | 2.30 |
| xrdev | -0.81 | 15.57 | -115.14 | 84.37 |

Table 2.3: Summary statistics of variables

# Chapter 3

# BMA Methodology

In this chapter we provide a description of the BMA method which will serve two purposes. First, it will present the BMA method in general for those who are interested in the method itself. Second, this description will pave the way towards the interpretation of results. To present the BMA method we rely heavily on Koop (2003),Feldkircher & Zeugner (2009) and Zeugner (2011).

## 3.1 Shift from OLS framework

For the purpose of this thesis, we begin with the Ordinary Least Squares (OLS) framework, using matrix notation for the standard linear model

$$y = \alpha + X\beta + \varepsilon, \tag{3.1}$$

where $y$ being the explained variable vector, $\alpha$ being the intercept vector, $X$ is the data matrix, $\beta$ stands for the matrix of slope coefficients, and the $\varepsilon$ being a vector of normal independent and identically distributed (iid) error terms, $\varepsilon \sim N(0, \sigma^2 I)$.

The elements of the Equation 3.1 have following dimensions, respectively

$$[n \times 1] = [n \times 1] + [n \times K][K \times 1] + [n \times 1], \tag{3.2}$$

where $n$ stands for the number of observations and $K$ is the total number of variables. The qualitative aspect of the regression imposes additional conditions on the matrices' properties. There are reasons for which both the feasibility and efficiency of this regression might be affected. The most important reason is the data matrix size. To fulfill the criteria of normality it is supposed that

$$n > 30. \tag{3.3}$$

And second, for the regression efficiency, it is required that

$$n > K. \tag{3.4}$$

Otherwise, the result of the regression could become inefficient or even unfeasible.

In practice, standard OLS regression consists of several iterative steps when the least statistically significant variables are eliminated and the regression is re-runed. By performing these steps, the data matrix X is shrinking and its size can be described as

$$[n \times (K - m)], \tag{3.5}$$

where $m$ stands for the number of eliminated variables. Usually, the common practice is either to have a strong theoretical support about which variables one should include and sticks to them or opposingly, to provide whichever statistically significant results. It is true that in special cases of known relationships between explained and explanatory variables, both of these options may occur. Unfortunately, very often the latter stated happens. Especially for the forecasting, when you need to explain a variable in terms of the others, having no prior knowledge about their relationships, the researcher faces the arbitrary decision which one he or she will include. The researcher then takes into account the most probable variables trying to eliminate the least statistically significant among them. As fewer and fewer variables are included, the original matrix $K$ shrinks by $m$, as described in the Equation 3.5. Which means, in other words, that $m$, the number of eliminated variables has to be big enough to assure p-value of each variable attaining an asterisked level (p-value $\leq 10\%$).

While iterating and eliminating insignificant variables, the researcher lacks two things. First of them is the lack of control over the true importance of the eliminated variables as the original matrix $K$ becomes smaller and smaller. The second problem is that there are many possible iteration paths leading towards different variables included and towards different results.

The BMA method attempt to control for both of these things. These two things might be together called the *model uncertainty* which is a commonly used name for the lack of knowledge about the inclusion of variables.

In following sections of this chapter, we show how the method works the model uncertainty out.

## 3.2 Submodel Structure

To show how the BMA works, it is appropriate to split into submodels, each of them using different set of variables. The data sub-matrix $X_i \in X$.

We have a model of data sub-matrix $X_i$, while taking into account the original full-size of the matrix $X$, with dimension $[n \times K]$, the $X_i$ matrix being of size $[n \times K_i]$.

$$y = \alpha_i + X_i\beta_i + \varepsilon, \tag{3.6}$$

where $i \in [1, 2^K]$ as described in the Section 3.11.

There are two reasons why to define a model with the intercept being aside from the vector of coefficients $\beta_i$. The first reason is that this model works even if the data are centered around zero. The other reason is to keep the intercept in the "game" when other variables from $X_i$ vary and in an extreme case, no variable is selected from the matrix $X_i$.

## 3.3 Bayes Theory

The aim of this section is to introduce the Bayesian point of view on the probability. This section is supposed to provide the analytical background of the BMA method for the estimates of this thesis.

Suppose $A$ and $B$ are two random variables with probabilities to occur *p(A)* and *p(B)*. The joint probability of these two events happenning together can be described by rules of probability as

$$p(A \cap B) = p(A|B)p(B) \tag{3.7}$$

and equally as

$$p(A \cap B) = p(B|A)p(A). \tag{3.8}$$

Together these two equations yield

$$p(A|B)p(B) = p(B|A)p(A), \tag{3.9}$$

which if rewritten form the Bayes theorem

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}. \tag{3.10}$$

A property the Bayesians believe is that the regression coefficients are not given and according to this point of view, $\beta$, the matrix of coefficients is a random variable following a probability distribution. According to the rule of conditional expectation, this can be written as

$$E(\beta|y, X) = p(\beta|y, X). \tag{3.11}$$

## 3.4 Posterior Mean

Substituting $A = \beta$ and $B = y, X$ into the Equation 3.10, where $\beta$ represents the coefficents and $y, X$ are the regression data, we have

$$p(\beta|y, X) = \frac{p(y, X|\beta)p(\beta)}{p(y, X)}, \tag{3.12}$$

where

- $p(\beta|y, X)$ is the posterior mean of the coefficient $\beta$, given the data,
- $p(y, X|\beta)$ is the Marginal Likelihood (ML) or the data-generating process given $\beta$,
- $p(\beta)$ is a prior density of the parameter $\beta$,
- $p(y, X)$ is the probability of the data.

All of these items will be further discussed in this section.

But back to the $\beta$'s posterior mean which is also equal to:

$$E(\beta|y, X) = \sum_{i=1}^{2^K} p(M_i|y, X)E(\beta_i|M_i, y, X), \tag{3.13}$$

where

- $p(M_i|y, X)$ is the Posterior Model Probability (PMP) of the $i - th$ sub-model, given the data,

- $E(\beta_i|M_i, y, X)$ is the estimate of $\beta_i$ in the given $i-th$ submodel for the data $y, X$.

The posterior mean is affected by the choice of parameter prior $g$ as discussed in the Section 3.10 (Zeugner 2011). Then, we have

$$E(\beta_i|M_i, y, X, g) = \frac{g}{1+g}\beta_i, \tag{3.14}$$

where $\beta_i$ stands for a standard OLS estimate.

## 3.5 Posterior Variance

While providing the notation for the BMA posterior variance of the estimated coefficients, again we rely on the Moral-Benito (2010a) definition

$$\begin{aligned} var(\beta|y,X) &= \sum_{i=1}^{2^K} \underbrace{p(M_i|y,X)}_{\text{PMP}} var(\beta_i|M_i,y,X) + \\ &+ \sum_{i=1}^{2^K} \underbrace{p(M_i|y,X)}_{\text{PMP}} [E(\beta_i|M_i,y,X) - E(\beta|y,X)]^2. \end{aligned} \tag{3.15}$$

This expression is using the PMPs as weights to calculate the BMA's $\beta$ as a linear combination of following items

- $var(\beta_i|M_i, y, X)$ is the variance of the $\beta_i$ in the $i-th$ submodel,
- $E(\beta_i|M_i, y, X)$ is the expected value of $\beta_i$ in the $i-th$ submodel,
- $E(\beta|y, X)$ is the posterior mean defined in the Section 3.4 and described by the Equation 3.13.

The parameter prior $g$ described in the Section 3.10 affects the $\beta_i$ variance in the following way (Zeugner 2011)

$$cov(\beta_i|M_i, y, X, g) = \frac{(y-\bar{y})'(y-\bar{y})}{n-3} \frac{g}{1+g} \left(1 - \frac{g}{1+g} R_i^2\right) (X_i' X_i)^{-1}, \tag{3.16}$$

where $\bar{y}$ stands for the vector mean, the $'$ stands for the transposition and $R_i^2$ stands for $i-th$ model $R^2$.

## 3.6 Posterior Model Probability

In the BMA framework, the PMP serves as the crucial tool assigning the submodels results's weights for the estimates averaging. It denotes how likely is the $i-th$ model $M_i$ to be before looking at the data. Following formula for the PMP arises from the Bayes theorem in the Equation 3.10 (Zeugner 2011).

$$p(M_i|y, X) = \frac{p(y|M_i, X)p(M_i)}{p(y|X)}, \tag{3.17}$$

where

- $p(y|M_i, X)$ stands for the ML discussed in Section 3.7,
- $p(M_i)$ is the model prior described in Subsection 3.10.2 and
- $p(y|X)$ denotes the integrated likelihood across all models or the probability of $y$, given $X$.

As Zeugner (2011) and many other BMA researchers assert, the integrated likelihood in the denominator from the Equation 3.17 can be expressed as

$$p(y|X) = \sum_{i=1}^{2^K} p(y|M_i, X)p(M_i) \tag{3.18}$$

and is a constant term across all submodels. Therefore, the PMP of the $i-th$ model $p(M_i|y, X)$ is directly proportional (signed as $\propto$) to the ML of the $i-th$ model times the $i-th$ model prior as follows

$$p(M_i|y, X) \propto p(y|M_i, X)p(M_i). \tag{3.19}$$

## 3.7 Marginal Likelihood

From the sections above, it is apparent that the ML plays a crucial role in the BMA method. According to the chap.1 of Koop (2003), the $i-th$ model ML equals

$$p(y|M_i, X) = \int_\beta p(y|\beta_i, M_i, X)p(\beta_i|M_i, X)d\beta_i. \tag{3.20}$$

While the Equation 3.20 is the easiest form to display, it is difficult to calculate. That is why, for example Feldkircher & Zeugner (2009) use following formula providing the background for the calculation of the $i-th$ model ML function

$$p(y|M_i, g) = \int_0^\infty \int_\beta p(y|\beta_i, \sigma_i^2, M_i)p(\beta_i, \sigma_i^2|g)d\beta d\sigma, \tag{3.21}$$

newly employing a prior $g$ which will be further discussed in the Section 3.10. Further, the Equation 3.21 is supposed to be proportional (signed with $\propto$) to the Equation 3.22

$$p(y|M_i, X, g) \propto (y-\bar{y})'(y-\bar{y})^{-\frac{n-1}{2}} \times (1+g)^{-\frac{K_i}{2}} \times (1-\frac{g}{1+g})^{-\frac{n-1}{2}}. \tag{3.22}$$

Equation 3.22 deploys the technique assigning the prior $g$ on both the model and parameter space, it accounts for the $i-th$ model size penalty for the number of included variables $K_i$, takes into account the $i-th$ model data variation $(y-\bar{y})$ and the number of observations $n$.

## 3.8 Posterior Inclusion Probability

The probability that a variable is included in the true model can be calculated according to Moral-Benito (2010a) as

$$p(\beta_k \neq 0|y, X) = \sum_{i=1}^{2^K} p(M_i|\beta_k \neq 0, y, X), \tag{3.23}$$

where $p(M_i|\beta_k \neq 0, y, X)$ stands for the PMP with the property that it is added up only in case that the $\beta$'s $k-th$ coefficient is different from zero.

## 3.9 Parameter Positivity

Another property provided as the result is the probability of one of the $\beta$'s coefficients being positive. The parameter positivity as shown in Koop (2003) can be rewritten as follows

$$p(\beta_k \geq 0|y, X) = \sum_{i=1}^{2^K} p(\beta_{i_k}|M_i, y, X)p(M_i|y, X). \tag{3.24}$$

Which means that this statistics simply creates a linear combination of the $\beta$'s $k-th$ coefficient weighted by PMP across all models.

## 3.10 Priors

To perform the BMA, two types of priors have to be specified. These priors bring an additional information about various distributions of parameters and model size. The first one is the parameter prior, the other the prior model size. Some authors prefer to insert no information on the submodels $M_i$ and the coefficients $B_i$ (Eicher *et al.* 2009). We show the purpose and the most popular setting in the Subsection 3.10.1 and Subsection 3.10.2 respectively, following Moral-Benito (2010b).

### 3.10.1 Parameter Priors

The function of the parameter priors is to determine the posterior mean of the slope coefficient $\beta_i$ arising as $p(\beta)$ from Equation 3.12. The hyperparameter $g$ reflects how certain the researcher is that the coefficients are zero (Zeugner 2011). Small $g$ stands for small variance and in the hypothesis testing framework affects the potential rejection of coefficients $\beta_i$ being zero more easily. As $g \to \infty$, the coefficient estimator approaches the OLS estimates.

Each model $M_i$ requires its own prior for $\beta_i$. Because of huge quantity of the candidate models, a simplification reducing prior-assigning time is among the most popular objective.

The most of the BMA literature favors the natural-conjugate approach, supposing normal distribution of the coefficient $\beta_i$, zero mean and the variance proposed by Zellner (1986) employing prior covariance given by $g(X_i'X_i)^{-1}$, where the prior $g$ assigns the importance to the researcher's beliefs. The conditional prior on the coefficient $\beta_i$ is as follows:

$$\beta_i|\sigma^2, M_i, g \sim N(0, \sigma^2 g(X_i'X_i)^{-1}). \tag{3.25}$$

Where the variance parameter $\sigma$ is common to all the models considered. Fernandez *et al.* (2001) proposes an uninformative prior $p(\sigma) \propto \sigma^{-1}$ and also sets the constant term prior $p(\alpha_i) \propto 1$. While the hyperparameter $g$ still remains to be assessed. Moral-Benito (2010b) lists three most popular settings for this hyperparameter $g$:

**(a)** *Unit Information Prior* (g-UIP), $g = n$,

**(b)** *Risk Inflation Criterion* (g-RIC), $g = K^2$,

**(c)** *Benchmark Prior* (g-BRIC), $g = max\{n, K^2\}$,

where $n$ and $K$ stands for the data matrix $X$ dimensions, being $[n \times K]$.

Fernandez *et al.* (2001) determined the *Benchmark prior*, combining the two previous settings, as the one with the best predictive performance. Therefore, in this paper we challenge the *Benchmark prior* results with other options in Chapter 6.

Among the alternative settings of the parameter prior $g$ belong so called hyper prior settings influencing the prior expected shrinkage factor $E(\frac{g}{1+g})$. Feldkircher & Zeugner (2009) propose a Benchmark Prior setting for $g$, guaranteeing asymptotic consistency as follows

$$E(\frac{g}{1+g}) = \frac{2}{a}, \tag{3.26}$$

where the parameter $a \in ]2, 4]$.

### 3.10.2 Model Priors

Last thing to determine for the BMA method is the model prior probability $p(M_i)$. In the BMA framework, this prior is used to assess the averaging weights represented by the posterior model probability (PMP) as arises from the Equation 3.17. Following Moral-Benito (2010b), the most common setting is the Binomial distribution representing the fact that a variable is either included in $i - th$ model $M_i$ or not.

This property affects the expected model size ($\Xi$) as follows

$$\Xi \sim Bin(K, \xi), \tag{3.27}$$

where $K$ is the total number of variables of the data matrix $X$ and $\xi$ is the prior inclusion probability assigned to each variable. Assuming any prior knowledge about variable inclusion meaning the researcher does not influence any model $M_i$ probability it is common to expect a fair fifty-fifty value

$$\xi = \frac{1}{2}. \tag{3.28}$$

Given the Equation 3.27 the prior model probability can be expressed as

$$P(M_i) = \xi^{K_i}(1-\xi)^{K-K_i}, \tag{3.29}$$

where $K$ is the total number of variables in the data matrix $X$ while $K_i$ is the submatrix of $X$ based on selected variables only.

Plugging the Equation 3.28 into the Equation 3.29 simplifies into

$$P(M_i) = \frac{1}{2^K}. \tag{3.30}$$

This means that under this setting the prior probability of various models $M_i$ is spread uniformely across the whole model space.

The property of binomial distribution yields that

$$E(\Xi) = K \times \xi, \tag{3.31}$$

where

- $E(\Xi)$ is the expected final model size,
- $K$ is one of the data matrix $X$ dimensions and is given,
- $\xi$ is the universal variable's prior inclusion probability.

This simple relationship shows the possibility of fixing different priors because only $K$ is given. For example, the chosen property in the Equation 3.28 yields the expected model size $E(\Xi) = \frac{K}{2}$. But it is also possible to set prior expected model size yielding an appropriate value for the prior inclusion probability. These properties are used to penalize the $i-th$ model $M_i$ for its size.

Among the most frequent prior settings one can let the $\xi$ to be drawn from a distribution instead of being fixed at a half. Manual ellicitation of the variables is less frequent.

## 3.11 Model space

Building a model which takes control over all variables in the $X$ matrix looks promising but it has some limits on its own. When a researcher faces a decision about including a variable, he has two possibilities: include or eliminate represented by the binomial distribution of the prior model probability as announced by Equation 3.27.

This logic of binomial distribution determines the model space also which starts at 2 and doubles with each additional variable. The model space ($\Psi$) is then

$$\Psi = 2^K. \tag{3.32}$$

The model candidate space $M$ is then represented by

$$M = \{M_1, M_2, \ldots, M_{2^K}\}. \tag{3.33}$$

The computational burden increases geometrically with additional variable while the number of observations $n$ has only a linear influence.

Having all possible combinations of the data subsamples $X_i$ to evaluate as submodels $M_i$, it takes rapidly increasing quantity of time. For example, if there was a hundred variables in the data matrix $X$, the dimension $K$ would be $K = 100$, determining the total model space $\Psi = 2^{100}$. If a computer is able to evaluate, let's say 1000 models per second, the time necessary to evaluate all possible combinations of submodels would still be $4 \times 10^{19}$ years. As illustrated by this example, the model space touches the computational limits as well as the time requirements.

The Section 3.12 discusses how the BMA method overcome this barrier.

## 3.12 MCMC Sampling

In case of a large number of variables $K$, the ellicitation of all models becomes infeasible due to number of possible covariate combinations. These combinations, referred as model space, cause a heavy computational burden and often does not allow to ellicit the true BMA results. The time requirements in the BMA method are substituted by a sampling algorithm visiting only part of combinations of variables $X_i$ to achieve the results in some reasonable and adequate time period (Feldkircher & Zeugner 2009).

### 3.12.1 Sampling Algorithm

The BMA method employed in this thesis uses an approximation. As the Bayesians believe that models are random, they employ sampling algorithms to draw from the model probability distribution. Zeugner (2011) provided a description of an Monte Carlo Markov Chain (MCMC) sampler for the BMA method, relying on the Metropolis-Hastings algorithm which is as follows.

*(i)* a random model $M_i$ called "current" with PMP $p(M_i|y, X)$ is drawn,

*(ii)* a candidate model $M_j$ is proposed,

*(iii)* the sampler switches to $M_j$ with probability

$$p_{i,j} = min\{1, \frac{p(M_j|y, X)}{p(M_i|y, X)}\}, \tag{3.34}$$

*(iv)* **(a)** if $M_j$ is accepted, it becomes the current model,

**(b)** if $M_j$ is rejected, the sampler draws another candidate model facing the current model $M_i$,

*(v)* whole procedure repeats.

Based on the Equation 3.34, this algorithm walks through the model space aiming to determine the models with the highest PMP. To limit the time required, using the "*bms*" package developed by Zeugner and Feldkircher, one can specify the number of iterations, the number of initial low-quality models that are omitted as well as the sampler settings which operate as follows:

**(a)** *Birth-death*: randomly chosen $K - th$ variable is

*(i)* added to the $X_i$ if not already included

*(ii)* removed from $X_i$ if included

**(b)** *Reversible-jump*

*(i)* with 50 % chance it behaves as (a)

*(ii)* with 50 % chance it drops one variable and adds another one

**(c)** *Enumeration* goes through all models $M_i$.

### 3.12.2 Algorithm Quality

There are two facts affecting the approximation quality. The MCMC sampler converges in number of iterations to the true BMA model. First, as the initial draws of the MCMC sampler statistically exerce lower PMP than the later draws, it is the number of models visited that influences the quality.

Second, complicated probability distributions of marginal likelihoods of the models impose difficulties for the sampler to converge towards the analytical values of the PMP (Zeugner 2011).

The quality of the algorithm sampling is measured by a correlation between analytical PMPs and the PMPs of the MCMC sampled models for an arbitrarily set quantity of the best models $M_i$ (with the highest PMPs). Usually, the number of best models is in thousands which on one hand contrasts with the total huge model space but on the other hand these best models can cover most of the posterior model probability mass. Which is the case when it is feasible to base posterior statistics on analytical likelihoods instead on the MCMC sampled ones (Zeugner 2011).

Levels of correlation above 0.99 indicate good convergence (Zeugner 2011). As there are two factors affecting the quality, the complicated likelihood distributions and the number of iterations, in case of smaller correlation coefficient, the number of MCMC iterations has to be enlarged.

# Chapter 4

# Model

The BMA method assumes several properties that are described in following paragraphs.

**Correlation of Variables** As a qualitative assumption, the BMA method requires low linear dependence among the variables. To address potential multicollinearity issues, we had to check these correlations first. Otherwise, the high correlations might influence the choice of variables across all models by offsetting the variables one against another in each particular model. For example, in case of two variables being highly correlated among themselves while none of these separately exerts any statistical significance when put together, they might be found significant both of them, having similar coefficents and just an opposite sign. Which would cause zero effect overall but the model would be spoiled by too many redundant, mutually offsetting variables. Therefore, the BMA requires a check of the correlation coefficients among explanatory variables. Prior to launch the research, several variables exhibiting values above 0.85 were eliminated. The correlation matrix shown in Table 4.1 includes explained variable — the *spread* — too.

**Spread Forecasting** To achieve the spread predicting power, we introduce one year lag into all of the explaining variables. This data lagging property is supposed to determine the spread size one year ahead. The names of variables were added an $L1$ prefix to symbolize the one year lag. The displayed correlation coefficients for the non-lagged variables (Table 4.1) are not supposed to vary importantly due to data stationarity which was ensured by the fixed effects transformation - the data demeaning.

**Fixed Effects** The panel data were demeaned by $j-th$ country units separately. This transformation turns the standard $OLS$ framework into fixed effects model framework. Introducing country fixed effects attempt to control for unobserved time-invariant country heterogeneity. Using fixed effects essentially means that we are using only time variation within countries to identify effects of variables on spreads.

Consequently, we estimate following equation:

$$y_{j,t} = X_{j,t-1}\beta + Z_j + \varepsilon_{jt}, \tag{4.1}$$

where

- $y_{jt}$ is the spread observed for the $j-th$ country at time $t$,
- $X_{jt}$ is a time-variant regressor for $j-th$ country at time $t$,
- $Z_j$ is a fixed effect of $j-th$ country,
- $\varepsilon_{jt}$ is an error term assumed to satisfy standard OLS assumptions.

| | spread | BEXP | DGDP | INVR | LRAT | NBTT | PSBR | PUDP | TDRA | UNEM | dlBEXL | dlCCPI | dlCEXP | dlFRES | dlLCHD | dlMIPD | dlRIND | dlXRPD | fresgdp | open | xrdev |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| spread | 1.00 | | | | | | | | | | | | | | | | | | | | |
| BEXP | -0.03 | 1.00 | | | | | | | | | | | | | | | | | | | |
| DGDP | 0.02 | -0.41 | 1.00 | | | | | | | | | | | | | | | | | | |
| INVR | -0.01 | -0.07 | 0.21 | 1.00 | | | | | | | | | | | | | | | | | |
| LRAT | -0.38 | -0.20 | 0.03 | -0.04 | 1.00 | | | | | | | | | | | | | | | | |
| NBTT | -0.08 | -0.10 | 0.05 | -0.07 | 0.21 | 1.00 | | | | | | | | | | | | | | | |
| PSBR | -0.08 | -0.16 | 0.29 | 0.18 | -0.06 | 0.19 | 1.00 | | | | | | | | | | | | | | |
| PUDP | 0.06 | 0.33 | -0.26 | -0.08 | -0.07 | -0.02 | -0.33 | 1.00 | | | | | | | | | | | | | |
| TDRA | 0.06 | -0.06 | 0.04 | 0.04 | -0.04 | 0.08 | 0.22 | 0.05 | 1.00 | | | | | | | | | | | | |
| UNEM | 0.16 | 0.13 | -0.03 | -0.07 | 0.25 | 0.06 | -0.37 | 0.08 | -0.27 | 1.00 | | | | | | | | | | | |
| dlBEXL | 0.01 | -0.31 | 0.33 | 0.03 | 0.26 | 0.20 | 0.03 | -0.30 | 0.01 | 0.02 | 1.00 | | | | | | | | | | |
| dlCCPI | 0.23 | -0.28 | 0.19 | 0.03 | 0.39 | 0.14 | 0.07 | -0.18 | -0.02 | 0.15 | 0.53 | 1.00 | | | | | | | | | |
| dlCEXP | -0.02 | -0.18 | 0.62 | 0.13 | 0.07 | -0.04 | 0.08 | -0.11 | 0.00 | 0.08 | 0.11 | 0.11 | 1.00 | | | | | | | | |
| dlFRES | 0.04 | -0.16 | 0.09 | 0.04 | 0.06 | 0.12 | 0.02 | -0.07 | 0.06 | 0.00 | 0.11 | 0.05 | -0.02 | 1.00 | | | | | | | |
| dlLCHD | -0.14 | -0.12 | 0.35 | 0.00 | -0.08 | 0.15 | 0.13 | -0.20 | -0.08 | -0.06 | 0.27 | 0.08 | 0.12 | 0.14 | 1.00 | | | | | | |
| dlMIPD | -0.13 | -0.14 | 0.52 | 0.17 | -0.02 | 0.08 | 0.23 | -0.13 | -0.03 | 0.01 | 0.12 | 0.06 | 0.43 | 0.01 | 0.38 | 1.00 | | | | | |
| dlRIND | 0.02 | -0.27 | 0.85 | 0.14 | 0.01 | 0.00 | 0.21 | -0.14 | 0.07 | 0.04 | 0.20 | 0.14 | 0.68 | 0.07 | 0.29 | 0.51 | 1.00 | | | | |
| dlXRPD | 0.23 | -0.05 | -0.17 | 0.03 | 0.20 | -0.03 | -0.08 | 0.04 | 0.06 | 0.11 | 0.05 | 0.30 | -0.04 | -0.07 | -0.80 | -0.36 | -0.17 | 1.00 | | | |
| fresgdp | 0.04 | -0.47 | 0.21 | 0.38 | -0.06 | 0.02 | 0.10 | -0.13 | 0.21 | -0.23 | 0.04 | -0.04 | 0.07 | 0.15 | 0.04 | 0.05 | 0.13 | -0.05 | 1.00 | | |
| open | 0.12 | -0.20 | 0.13 | 0.58 | -0.19 | -0.17 | 0.08 | -0.04 | 0.25 | -0.17 | -0.04 | -0.10 | 0.07 | -0.01 | -0.01 | 0.09 | 0.05 | -0.05 | 0.76 | 1.00 | |
| xrdev | -0.03 | -0.12 | 0.08 | -0.13 | -0.06 | 0.08 | -0.02 | -0.11 | -0.02 | -0.09 | 0.13 | 0.02 | -0.09 | 0.01 | 0.26 | 0.03 | -0.01 | -0.24 | 0.10 | -0.04 | 1.00 |

Table 4.1: Correlation matrix of variables

# Chapter 5

# Results

This chapter is consists of three sections. First part provide the model diagnostics. Second part shows the BMA estimates. Third part is dedicated to a comparison of own findings with the relevant literature in two aspects: in an inclusion of selected variables in the true model and in a direction of an effect.

## 5.1 Model Diagnostics

This section provides a summary of the model diagnostics. It comprises four items. First, the Table 5.1 describes the results as well as the initial priors entered in the BMA analysis. The Figure 5.1 and the Figure 5.2 represent graphically the MCMC algorithm approximation quality while the former stated figure is more general, the latter more detailed. The Figure 5.3 plots the prior and the posterior model size densities.

| | | |
|---|---|---|
| Mean no. regressors | Draws | Burnins |
| "12.5540" | "4e+06" | "2e+06" |
| Time | No. models visited | Modelspace 2^K |
| "1.034798 hours" | "2369097" | "1e+06" |
| % visited | % Topmodels | Corr PMP |
| "226" | "94" | "0.9995" |
| No. Obs. | Model Prior | g-Prior |
| "616" | "random / 10" | "hyper (a=2.003247)" |
| Shrinkage-Stats | | |
| "Av=0.9604, Stdev=0.018" | | |

Table 5.1: Model Summary

**Diagnostic Vocabulary** Among the most important diagnostics tools belong the

**(a)** *PMP Correlation* represents the quality, anything above 0.99 is a good result (Subsection 3.12.2),

**(b)** *% Topmodels* displays the sum of PMPs of best 5000 models,

**(c)** *% visited* is a share of evaluated models with respect to the total model space,

**(d)** *g-Prior* describes the $g$ parameter settings described in the Subsection 3.10.1,

**(e)** *Model Prior* is the way of elliciting the prior model size followed by its expected value $\frac{K}{2}$.

The Figure 5.1 and Figure 5.2 depict the quality of the results gathered by MCMC sampling. Both of them show the correlation between analytical and sampled PMPs. The former is based on 5000 models with the highest PMPs. The latter is supposed to demonstrate the detailed view and is based on best 30 models. The purpose of the detailed version is twofold. First, to show that there are two lines and second, that these lines are not decreasing purely exponentionally. The levels of correlation shown, reach the same value of 0.9995 in both charts. As this value approaches very closely to unity, the consequenting quality of the MCMC sampling tends to be satisfactorily high. Partly, this is due to the fact that the number of iterations is twice as high as the model space. Which means that the setting was approximating the enumeration of all the possible models $M_i$ while the element of randomness was kept alive for the illustration of the BMA method MCMC sampling in case of many variables.

There is no surprise that the MCMC sampled line is refracted as the single models $M_i$ are drawn. What is surprising is the shape of the analytical curve which is smooth in the Figure 5.1 and contrasts with being refracted as depicted in the more detailed version in the Figure 5.2. This might be explained by the complicated marginal likelihoods of the single $i - th$ models $M_i$.

The Figure 5.3 shows the prior and the posterior model size densities. The uniform prior distribution having expected value in the middle turns out by the calculation process into normally distributed posterior density, slightly skewed positively and slightly shifted positively as well (towards larger number of explanatory variables). Finally, the mode of the posterior expected model size is at twelve regressors, with the mean determined in the Table 5.1 as 12.554 variables. This means that the true model describing the spread and using our set of variables would most likely contain twelve or thirteen determinants. The selection of the twelve or thirteen best spread describing variables from our dataset is handled in sections that follow.

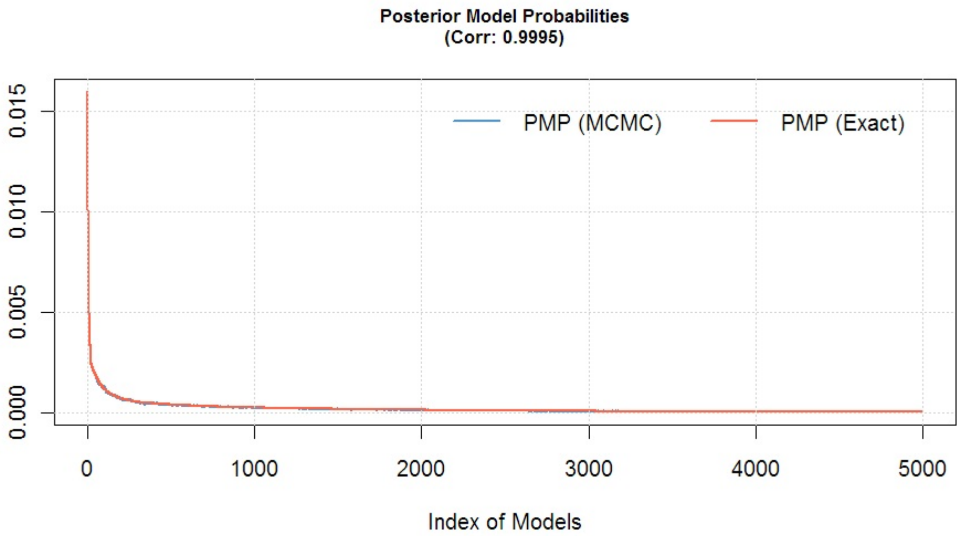


Figure 5.1: PMP Convergence of 5000 Best Models

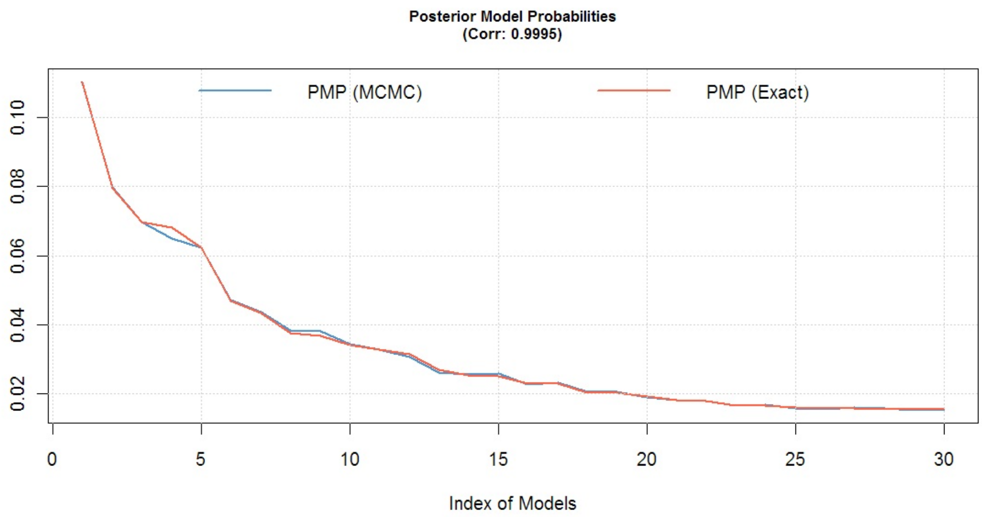


Figure 5.2: PMP Convergence of 30 Best Models

## 5.2 Model Inclusion

To be able to determine the best twelve or thirteen variables from our dataset most likely to explain the spread we use the Figure 5.4. This figure is a graphic

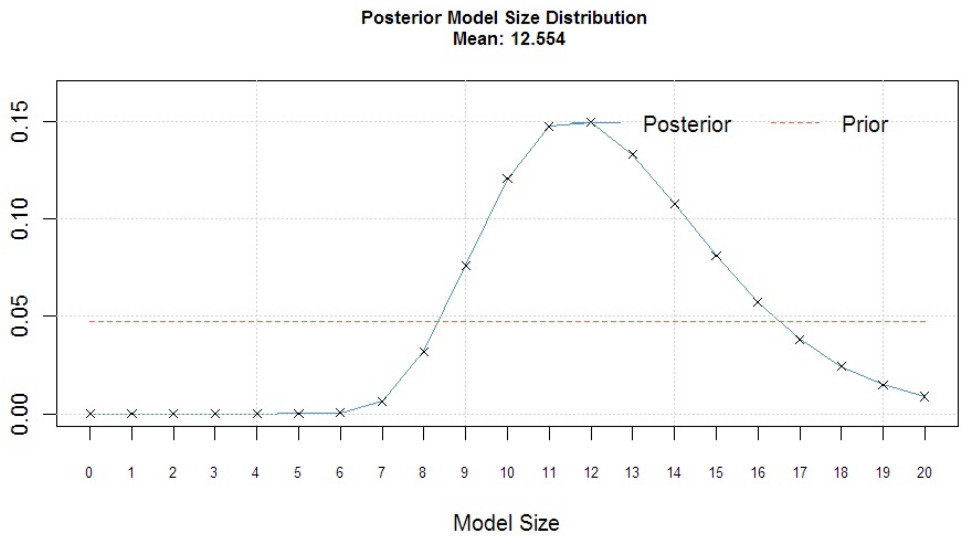


Figure 5.3: Posterior Model Size

representation of the BMA results. A description how this figure was built follows.

The Figure 5.4 is based on 5000 models $M_i$ with the highest PMPs (the PMP is described in the Section 3.6). The x-axis cumulative PMP sums up to 1 over the whole model space. Even though the total model space is large, a small fraction of them covers most of the model probability. Based on the Equation 3.32 and the number of variables $K$ in the data matrix $X$ being $K = 20$ is 1048576. In the Figure 5.4 we take into account only the best 5000 models which is less than 5% of the total model space, the cumulative PMP reaches 0.94.

The Figure 5.4 is a graphical summary of the BMA results, while the coefficients' sizes remain disregarded yet. We provide a description of the Figure 5.4 in the next paragraph.

The single submodels $M_i$ are organised in columns by their PMPs in descending order. Variables $K$ are in rows. The blank space in the column means that the $K-th$ variable is not included in the particular $i-th$ model $M_i$. Coloured field stands for the variable inclusion in the model $M_i$, while red stands for a negative effect in the $i-th$ model and blue for the positive sign. In the framework of the spread determinants, the negative effect of a variable in $i-th$ model $M_i$, marked by red color, stands for closing the spread while blue color represent spread-widening variables. The column width represents the PMP of model $M_i$. On the x-axis, the PMP of the submodels $M_i$ cumulates.

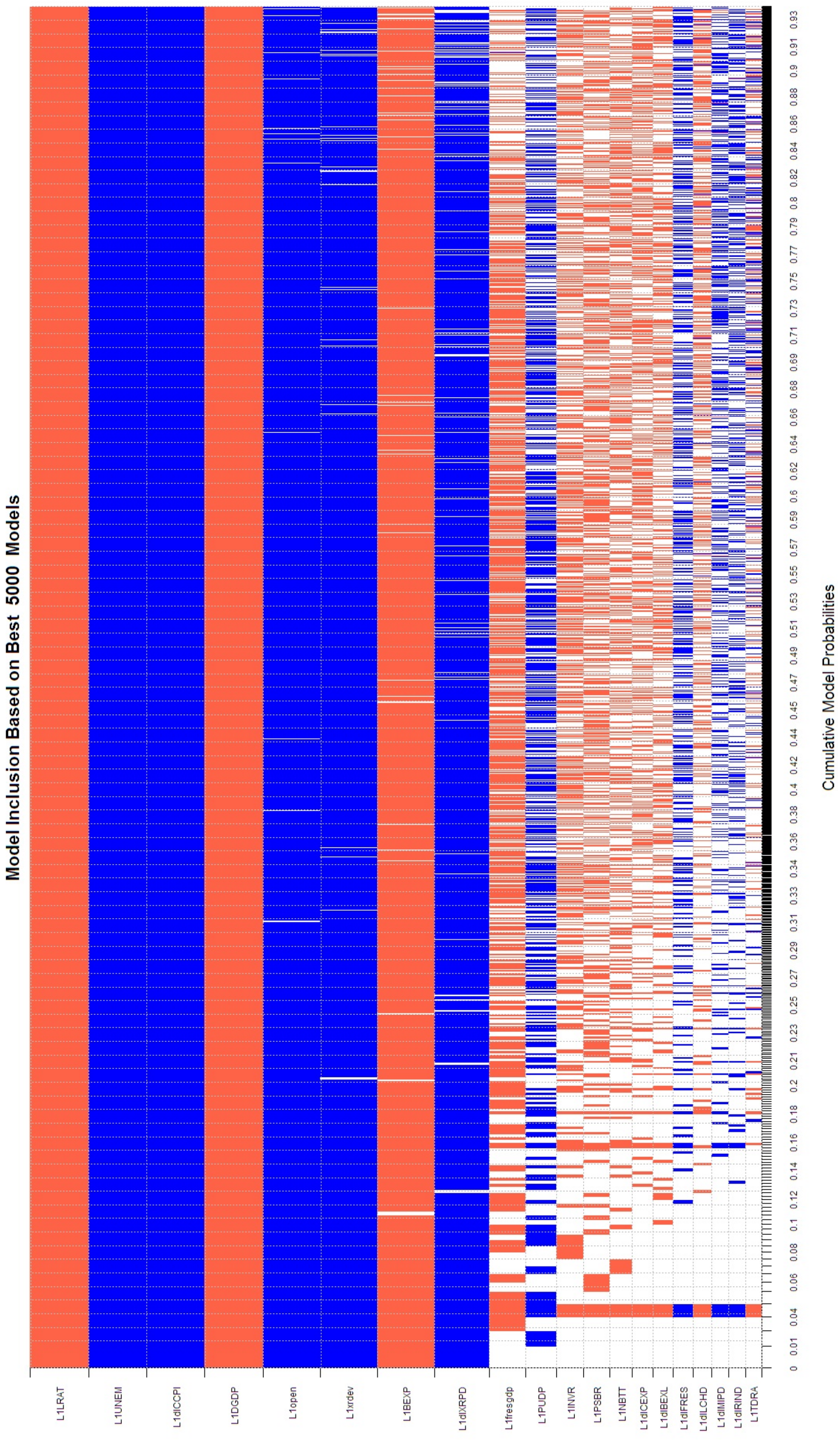


Figure 5.4: Model Inclusion

The row height represents the $K-th$ variable PIP, the construction of the PIP is described in the Section 3.8.

The visual pattern in the Figure 5.4 amplifies the posterior model design. The biggest coloured areas represent the explanatory variables most likely to explain the dependent variable, the spread. As the coloured fields shrinks in size and in density, the importance of a variable's explaining power declines too.

The model uncertainty problematics of the standard modelling aproach is well illustrated by model $M_i$, fourth in row according the highest PMP score which can be seen in the Figure 5.4. In this particular case, the researcher would face an unpleasant truth that sometimes all the included dataset variables are statistically significant even though their explaining power in terms of the PMP is limited as a whole. The researcher's next step would consist of a random exclusion of one or several variables trying to determine which of them might be among the best descriptors. Statistical significance might be a good reason to publish but a curious researcher would go further and would try to eliminate one or more variables. Followed by a surprise that the significance is lost for many other determinants, he would try another combination of determinants yielding similar results. If this researcher went through all the possible combinations and averaged the results, he would have done a similar thing as the "enumeration" algorithm used in the BMA method. Therefore, if he really was that curious (and uncertain), he would become a Bayesian.

## 5.3 Model Estimates

While the Section 5.2 was oriented on the graphical representation of the results, this section provides the coefficients, averaged and weighted by their importances (PMPs). The Table 5.2 shows the results while the variables are ordered by the posterior inclusion probability (PIP).

Using the Table 5.2 we are able to select the best twelve or thirteen[1] spread-explaining variables from our dataset.

Now, as we have determined the most likely spread-explaining variables, we will check the relevant literature results and compare the inclusion of the variables.

[1]The number being reported by the BMA method in the Table 5.1

| Variable | Variable Description | PIP | Post Mean | Post SD | Pos. | Idx |
|---|---|---|---|---|---|---|
| L1LRAT | lending interest rate | 1.000 | -0.148 | 0.012 | 0.00 | 4 |
| L1UNEM | unemployment | 1.000 | 0.159 | 0.028 | 1.00 | 9 |
| L1dlCCPI | inflation | 1.000 | 20.709 | 2.965 | 1.00 | 11 |
| L1DGDP | real GDP growth | 1.000 | -0.169 | 0.044 | 0.00 | 2 |
| L1open | openness ratio $(\frac{X+M}{2GDP})$ | 0.992 | 1.339 | 0.543 | 1.00 | 19 |
| L1xrdev | exchange rate deviation from trend | 0.982 | 0.026 | 0.009 | 1.00 | 20 |
| L1BEXP | budget expenditure / GDP | 0.974 | -0.027 | 0.010 | 0.00 | 1 |
| L1dlXRPD | exchange rate (domestic/USD) change | 0.947 | 4.448 | 1.713 | 1.00 | 17 |
| L1fresgdp | foreign exchange reserves (excl.gold) / GDP | 0.621 | -0.946 | 0.989 | 0.00 | 18 |
| L1PUDP | public debt / GDP | 0.537 | 0.003 | 0.003 | 1.00 | 7 |
| L1INVR | foreign direct investment / GDP | 0.460 | -0.014 | 0.022 | 0.00 | 3 |
| L1PSBR | budget balance / GDP | 0.451 | -0.016 | 0.026 | 0.00 | 6 |
| L1NBTT | terms of trade | 0.391 | -0.002 | 0.004 | 0.00 | 5 |
| L1dlCEXP | exports growth | 0.352 | -0.535 | 1.187 | 0.00 | 12 |
| L1dlBEXL | budget expenditure growth | 0.345 | -0.435 | 1.028 | 0.00 | 10 |
| L1dlFRES | foreign exchange reserves (excl.gold) growth | 0.343 | 0.093 | 0.216 | 1.00 | 13 |
| L1dlLCHD | labor cost growth | 0.323 | -0.308 | 1.172 | 0.03 | 14 |
| L1dlMIPD | abroad debt service growth | 0.294 | 0.077 | 0.288 | 0.99 | 15 |
| L1dlRIND | real industry growth | 0.292 | 0.514 | 2.053 | 1.00 | 16 |
| L1TDRA | trade balance / GDP | 0.270 | 0.000 | 0.006 | 0.45 | 8 |

- *Note 1:* All variables were 1 year lagged and are marked by "L1" in their names.
- *Note 2:* "Pos." stands for a conditional probability of the effect sign being positive.

Table 5.2: Coefficient Estimates

## 5.4 Inclusion of Variables: A Comparison With The Literature

This section provides the Table 5.3 which aims at comparing the subsets of variables included in particular models. We use Rowland & Torres (2004) and Frankel & Saravelos (2010) as the benchmark for the comparison of the spread determinants. The former study comprises a third benchmark made by Cantor & Packer (1996). All of them studied spread, creditworthiness and rating while we have included a mix of the three sets of the explanatory variables.

To increase the number of the benchmark variables, we used several conventions turning the rating determinants into spread determinants. These conventions are as follows. The direction of the spread explaining variables' effect is opposite to the effect both on the creditworthiness (Rowland & Torres 2004) and rating (Kamin & von Kleist 1999; McGuire & Schrijvers 2003; Uribe & Yue 2006). Generally, this is caused by a simple fact that in normal times the lower spread stands for these two things, higher creditworthiness and better sovereign rating and moderate market sentiments (Eichengreen & Mody 1998).

| Variable | Variable Description | PIP | Effect | R&T | C&P | F&S |
|---|---|---|---|---|---|---|
| L1LRAT | lending interest rate | 1.000 | (-) | | | 13 |
| L1UNEM | unemployment | 1.000 | (+) | | | |
| L1dlCCPI | inflation | 1.000 | (+) | (+) | (+) | 15 |
| L1DGDP | real GDP growth | 1.000 | (-) | (-) | (-) | 12 |
| L1open | openness ratio ($\frac{X+M}{2GDP}$) | 0.992 | (+) | | | |
| L1xrdev | exchange rate deviation from trend | 0.982 | (+) | | | 24 |
| L1BEXP | budget expenditure / GDP | 0.974 | (-) | | | 5 |
| L1dlXRPD | exchange rate (domestic/USD) change | 0.947 | (+) | | | 24 |
| L1fresgdp | foreign exchange reserves (excl.gold) / GDP | 0.621 | (-) | (-) | | 25 |
| L1PUDP | public debt / GDP | 0.537 | (+) | (+) | | 22 |
| L1INVR | foreign direct investment / GDP | 0.460 | (-) | | | 3 |
| L1PSBR | budget balance / GDP | 0.451 | (-) | | | 9 |
| L1NBTT | terms of trade | 0.391 | (-) | | | 9 |
| L1dlCEXP | exports growth | 0.352 | (-) | | | 17 |
| L1dlBEXL | budget expenditure growth | 0.345 | (-) | | | 5 |
| L1dlFRES | foreign exchange reserves (excl.gold) growth | 0.343 | (+) | | | 25 |
| L1dlLCHD | labor cost growth | 0.323 | (-) | | | |
| L1dlMIPD | abroad debt service growth | 0.294 | (+) | | | |
| L1dlRIND | real industry growth | 0.292 | (+) | | | |
| L1TDRA | trade balance / GDP | 0.270 | (-) | | | 11 |
| n/a | GDP per capita | n/a | | | (-) | 12 |
| n/a | exports / GDP | n/a | | (-) | | |
| n/a | debt service / GDP | n/a | | (+) | | |
| n/a | external debt | n/a | | | (+) | 3 |
| n/a | debt / exports | n/a | | (+) | | |
| n/a | current account / GDP | n/a | | | | 11 |
| n/a | default dummy | n/a | | (+) | | |
| n/a | economic development indicator | n/a | | | (-) | |
| n/a | default history indicator | n/a | | | (+) | |
| n/a | money supply | n/a | | | | 19 |
| n/a | equity returns | n/a | | | | 13 |
| n/a | debt composition | n/a | | | | 10 |
| n/a | contagion | n/a | | | | 6 |
| n/a | political / legal | n/a | | | | 6 |

Table 5.3: Inclusion Comparison

- *Note 0*: R&T is Rowland & Torres (2004); C&P is Cantor & Packer (1996); F&S is Frankel & Saravelos (2010);
- *Note 1*: Our own spread explaining variables are ordered by PIP, Spread = (10yr sov debt yield - 3M IBOR);
- *Note 2*: Sign (+) or (-) represent the observed effect on spread (or negatively inversed if those were for creditworthiness or rating) in particular study, the number in the F&S column represents the number of occurrences observed in their overview;
- *Note 3*: The occurrence count of the variables from other studies facing one or more our variables were modified proportionately as follows: F&S category GDP split equally into L1DGDP and (GDP per capita), F&S category Reserves split equally between L1dlFRES and L1fresgdp, F&S category Real Exchange Rate split equally between L1dlXRPD and L1xrdev, F&S category Credit dealt as L1PUDP, F&S Category Current account was split equally into L1TDRA and (current account / GDP), F&S Category Exports or Imports dealt as L1dlCEXP, F&S Category Capital Flows dealt as L1INVR.

The overview made by Frankel & Saravelos (2010) was close to our findings in two things. First, the variables we included in our dataset cover the mass of their overview. Second, our variables, as ordered by PIP, represent more or less the descending quantity of inclusions they reported.

However, there are still few variables that were not included in either study. From one side, the variables that we omitted are listed on the bottom of the Table 5.3. The reasons to prevent these variables from inclusion into our dataset were as follows:

*(i)* These variables were included solitarily in the benchmark studies and do not exhibit much of the inclusion consistency. Which is represented by low or no inclusion number in the overview made by Frankel & Saravelos (2010), or marked by a solitary effect sign in just one of the other studies.

*(ii)* These variables are hard to find or require an arbitrary decision about a threshold.

*(iii)* These variables are highly correlated with other variables.

From the other side, we have included variables that were not included in none of the benchmark studies. A part of these variables scored low in the PIP and were not found likely to be included in the model. But another part of them ranked high by the PIP. As an example, we can mention the *unemployment* and the *openness ratio* in the form $\frac{X+M}{2GDP}$.

The BMA method also attributes relative importances to the variables from the Frankel & Saravelos (2010) overview. These importances create a rank of variables in the form of the PIP. The ellicitation of the probability of a variable inclusion opens a possibility of being measured and compared against each other. Compared to the overview which only summarized the variables based on the number of published studies suffering from the possible publication bias.

## 5.5 Signs of Coefficients: A Comparison With The Literature

In this part we compare own estimated effects with the literature, based on the coefficient signs. For the purpose of a comparison we use the Table 5.3.

Negative effect on the spread stands for closing the spread which is understood as a good condition of the economy, connotated with possible better sovereign rating and higher creditworthiness as the explanatory variable rises.

On the other hand, a positive sign represent a spread widening effect as the explanatory variable rises. Which is a bad signal of the economy connotated with worsening of ratings.

There can be a gap identified in the PIP column in the Table 5.2, graphically shown by an increasing quantity of blank space in the Figure 5.4. This gap separes first eight best spread-explaining variables from the rest. Our description will focus on these eight variables plus those whose inclusion effect is comparable as shown in the Table 5.3 instead of twelve or thirteen as provided by the expected model size.

**Inflation** As one of the most important spread determinants, the positive sign for the *inflation* seems to be intuitive. In case of inflation being relatively high, potential investors want to offset the inflation costs of their borrowings by charging higher long term interest rates. These long term interest rates form directly a part of the *spread* definition which explains the positive sign of the effect.

**Real GDP growth** *Real GDP growth* ranks high by the BMA method among other variables. At the same time this variable was found significant in both of the two studies and the overview mentions its inclusion twelve times. These two studies confirm the negative sign of the effect on the spread that we observe. The *GDP growth* show the speed of change of an economy and can be also seen as a productivity increase which explains the spread decrease as country grows in the *GDP*. The other explanation channel goes through the *GDP* that is used as weights for other determinants. If the *GDP* grows, ceteris paribus, the variables measured as shares on *GDP* shrinks. Which is for a certain group of variables a good signal, for example, especially for the public debt as a policy instrument.

**Foreign Exchange Reserves** *Foreign exchange reserves / GDP* is kind of an international "hostage". The more reserves a country has, the higher the liquidity is and the more chances against market volatility the central banks have (Eichengreen October 1995). Lower volatility goes hand in hand with higher credibility and better ratings consequenting in lower spread. Therefore, the identified effect of this variable on the spread is negative — spread narrowing. This story is supported by the evidence of own research, by Rowland & Torres (2004) and also by the Frankel & Saravelos (2010) overview mentioning it twenty-five times. Another measure of the foreign exchange reserves was took in the form of *Foreign exchange reserves growth*. This flow measure has only half the chance of inclusion than the stock measure (share on the GDP) which means the former is not as important. Contrarily to the stock measure, the effect of an increase of the reserves' growth speed on the spread is positive, spread widening.

**Indebtedness** We use *Public debt / GDP* as a measure of a country indebtedness. It is no surprise that the more a country becomes indebted, the higher is the risk and associated costs for its potential investors due to the obvious policy path as perceived by them. We stress the validity of results being limited by the linear modeling approach and the dataset property. This results are valid on certain range only. A debt can be a part of rational intertemporal choice but the political temptation of never having to repay the debt plays its role. The higher potential costs, as perceived by investors, increase the lending interest rate which is directly linked to the spread increase as the debt rises. Therefore, the positive sign confirmed by Rowland & Torres (2004) and supported by twenty-two inclusions in the overview study of Frankel & Saravelos (2010).

**Government Expenditure** The policy influence takes its place here. The government expenditure represented by *Budget expenditure / GDP* ranks high by its PIP. But the effect of *government expenditure* lacks a solid explanation. Its effect is negative - an increase in the government size will narrow the spread which is counter-intuitive because the rational political motivation would be to increase the debt yet further. It might work like a leverage - the bigger the government size is, the more space for controlling the debt it has. The second measure of the government spending is the *Budget balance / GDP*. The results show again a counter-intuitive result: negative effect on the spread but a minor importance suggesting the relative government spending size is more important spread determinant than the relative budget deficit size. Third in row and again with a negative - spread narrowing effect, the *budget expenditure growth* is with a rather small PIP closing this group of variables. The speed of change of the government expenditure does play only a minor role in determining the spread size. This paragraph shows an inclusion probability pattern that favorizes relative measures to the speed of the debt change.

**Exports and openness** The effect of relative size of exports to the GDP on the spread remains unclear. No report in the Frankel & Saravelos (2010) suggests low importance. According to Rowland & Torres (2004), the effect of *openness* in the form of $Export/GDP$ on the spread is negative. This would stand for the metaphorical "international hostage" lowering the spread as the export rise. But in our case, the *openness* notation as the $(\frac{Export+Import}{2GDP})$ yields a positive — spread widening — effect. We consider this enlargement of the notation to cause no change of the effect as the trade balance tends to be zero on average. The effect was supposed to be equivalent to that of the short notation. Nevertheless, the enlarged notation for the *openness* yields a very high PIP of 0.992 and a positive sign. The positive sign might be explained by possible volatility sourcing from the international demand exposition. The international demand exposition is like a double-edged sword, it can increase the country's competitiveness if the exports are rising or high enough. On the other hand if the export shrinks in spite of a sudden international demand decrease while the imports remain the same, the trade balance becomes negative. To equalize the current account, an outflow pumps the money out of the relevant economy. The most important way how this outflow takes place is via issuance of an external debt that increases indebtedness which influences the spread positively.

**Unemployment** The *unemployment* is one of those variables that were not included in any other benchmark study. However, the BMA method it considers as one of the most important descriptors of the future spread, assigning it by a PIP of 1.000 which says that the inclusion into the tru model is a sure thing. The direction of the effect is intuitive. The higher the *unemployment* is, the higher the output gap tend to be. The output gap is an inefficiency. This inefficiency is partly offset by borrowing the money and creating a debt to be able to redistribute. The effect of the indebtedness takes place. The indebtedness is described in this section too.

**Exchange rate** We use two variables for the representation of the exchange rate influence on the spread. One of them is own construction of an *Exchange rate trend deviation*. The other is an *Exchange rate change* measured towards the $USD$. The effect of the former stated on the spread is found positive which means that an overvaluation of the currency has a spread-widening effect. The explanation might go through the worsening conditions for the exports. The shrinking exports are causing a need to cover the trade deficit by issuing a debt. This explanation follows the logic of both the previous paragraphs on the exports and indebtedness. The *exchange rate change*, seen as a central bank instrument to limit the market volatility, exhibit positive effect, similarly to the formerly stated variable. Even though we have splitted the total number of inclusions equally between these two similar variables, the evidence of an inclusion is supported by a relatively high numbers in the Frankel & Saravelos (2010) overview. While these variables sound similarly, based on the low correlation coefficient of $-0.24$ in the Table 4.1 we kept both of them in the dataset.

**Interest rate** As the *interest rate* tends to be closely correlated with the interbank offered rates which form a part of the spread definition, it seems they play a not surprising role in its explanation. Another central bank tool to defend against market volatility, the *interest rate* ranks on the top above all the other variables. Its estimated effect is negative. The higher the *interest rate* is, the narrower will the spread be.

**Remaining variables** The variables ranking low by the PIP are either the speed of change variables which turned out to be bad spread predictors compared to their stock versions. The variables representing this group are *budget expenditure growth* and *foreign exchange reserves growth*. Or, the other group exercing lower PIPs are the real economy variables as for example the *Terms of trade*, the *trade balance / GDP*, the *exports growth*, the *real industry growth*, the *labor cost growth* and *abroad debt service growth*.

# Chapter 6

# Robustness check

This section aims to provide a sensitivity analysis for alternative BMA settings that might play a role for the magnitude and the sign of results as the literature suggests (Fernandez *et al.* 2001; Eicher *et al.* 2011). We use the priors as developed in the $R$ package by Zeugner (2011).

The benchmark results shown in the Chapter 5 are based on BMA method estimation using following items:

- *model size prior* = *random*,
- $g - prior = hyper - BRIC$, which stands for $g = max\{n, K^2\}$ while the *hyper* prefix is guaranteeing asymptotic consistency at the same time.

We will check our results' consistency in two alternative directions. First, by elliciting different parameter priors $g\{UIP, RIC\}$ respectively, everything else remaining the same. Because in our case $n > K^2$, the third possible alternative $g = BRIC$ is equal to the $g = UIP$ and will not be handled separately. Second, we test the results' consistency by setting different model size priors for *model size prior*=$\{$*uniform,fixed*$\}$. In the final part of this chapter presents a sensitivity analysis comprising these four BMA alternative results.

Each of the alternatives comprises a paragraph of comments and following four items:

1. diagnostics table,
2. prior and posterior model sizes plot,
3. model inclusion chart and
4. coefficients table.

## 6.1 Parameter prior sensitivity

### 6.1.1 g=UIP(BRIC)

In this section, we handle two alternative settings together. The $g = UIP$ setting stands for $g = n$ as described in detail in the Subsection 3.10.1, where $n$ is the number of observations or one of the data matrix $X$ dimensions. Another alternative setting $g = BRIC$ uses $g = max\{n, K^2\}$. Since the number of observations $n$, $n = 616$ and the number of variables $K$, $K = 20$, $g = max\{616, 400\} \Rightarrow g = 616$ and $g = n$ which is identical for both cases. That is why we inspect these two settings together as $g = UIP(BRIC)$.

What is striking on the first sight is the low *number of models visited* compared to the benchmark settings. Despite the number being close to one third of the benchmark count, the *PMP correlation* is surprisingly high and attains 0.9999.

The model size tends to be substantially lower, the *Mean number of regressors* in the Table 6.1 shows 8.2332 compared to the 12.5540 from the benchmark. The uniformely distributed model prior size turns out to be rather normally distributed as depicted in the Figure 6.1.

Following the Figure 6.2, the model inclusion amplifies the relative importance of the same eight, most likely the spread-explaining, variables from the benchmark in comparison with the remaining variables.

The coefficients are shown in the Table 6.2.

| | | |
|---|---|---|
| Mean no. regressors | Draws | Burnins |
| "8.2332" | "4e+06" | "2e+06" |
| Time | No. models visited | Modelspace 2^K |
| "34.02092 mins" | "841014" | "1e+06" |
| % visited | % Topmodels | Corr PMP |
| "80" | "100" | "0.9999" |
| No. Obs. | Model Prior | g-Prior |
| "616" | "random / 10" | "UIP" |
| Shrinkage-Stats | | |
| "Av=0.9984" | | |

Table 6.1: g=UIP(BRIC) Summary

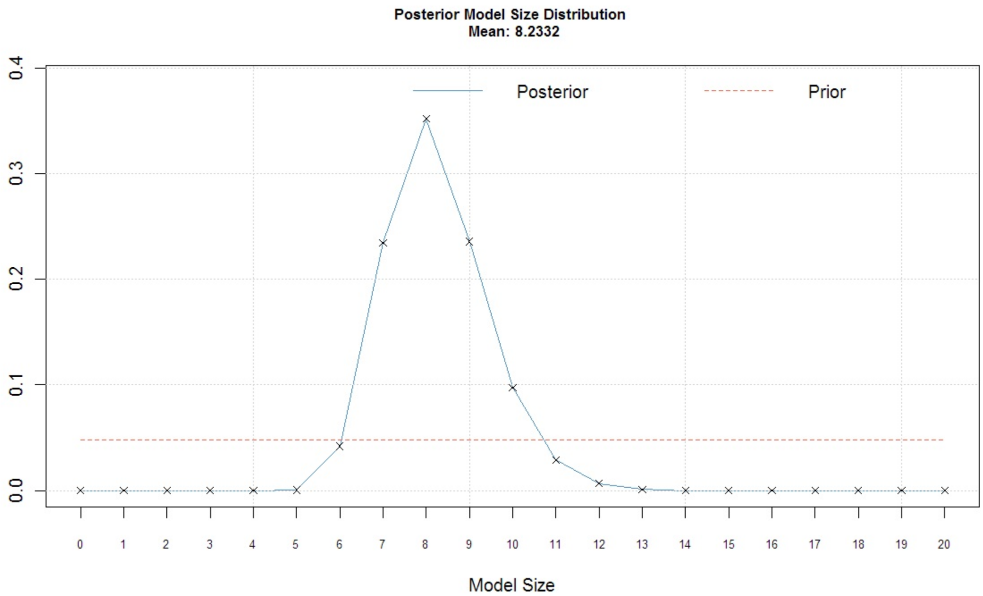


Figure 6.1: g=UIP(BRIC) Model Size

| Variable | Variable Description | PIP | Post Mean | Post SD | Pos. | Idx |
|---|---|---|---|---|---|---|
| L1LRAT | lending interest rate | 1.000 | -0.156 | 0.013 | 0.000 | 4 |
| L1UNEM | unemployment | 1.000 | 0.170 | 0.027 | 1.000 | 9 |
| L1dlCCPI | inflation | 1.000 | 21.687 | 2.983 | 1.000 | 11 |
| L1DGDP | real GDP growth | 0.996 | -0.179 | 0.037 | 0.000 | 2 |
| L1dlXRPD | exchange rate (domestic/USD) change | 0.893 | 4.358 | 1.835 | 1.000 | 17 |
| L1open | openness ratio ($\frac{X+M}{2GDP}$) | 0.877 | 0.901 | 0.483 | 1.000 | 19 |
| L1xrdev | exchange rate deviation from trend | 0.801 | 0.022 | 0.013 | 1.000 | 20 |
| L1BEXP | budget expenditure / GDP | 0.755 | -0.020 | 0.014 | 0.000 | 1 |
| L1dlLCHD | labor cost growth | 0.137 | -0.402 | 1.172 | 0.001 | 14 |
| L1PUDP | public debt / GDP | 0.128 | 0.001 | 0.002 | 1.000 | 7 |
| L1fresgdp | foreign exchange reserves (excl.gold) / GDP | 0.117 | -0.139 | 0.538 | 0.101 | 18 |
| L1PSBR | budget balance / GDP | 0.108 | -0.005 | 0.016 | 0.000 | 6 |
| L1INVR | foreign direct investment / GDP | 0.098 | -0.003 | 0.013 | 0.036 | 3 |
| L1NBTT | terms of trade | 0.079 | 0.000 | 0.002 | 0.000 | 5 |
| L1dlBEXL | budget expenditure growth | 0.047 | -0.062 | 0.415 | 0.000 | 10 |
| L1dlCEXP | exports growth | 0.045 | -0.063 | 0.454 | 0.000 | 12 |
| L1dlFRES | foreign exchange reserves (excl.gold) growth | 0.038 | 0.007 | 0.069 | 1.000 | 13 |
| L1dlRIND | real industry growth | 0.037 | 0.023 | 0.848 | 0.827 | 16 |
| L1dlMIPD | abroad debt service growth | 0.033 | 0.006 | 0.095 | 0.918 | 15 |
| L1TDRA | trade balance / GDP | 0.032 | 0.000 | 0.002 | 0.586 | 8 |

- *Note 1:* All variables were 1 year lagged and are marked by "L1" in their names.
- *Note 2:* "Pos." stands for a conditional probability of the effect sign being positive.

Table 6.2: g=UIP(BRIC) Coefficient Estimates

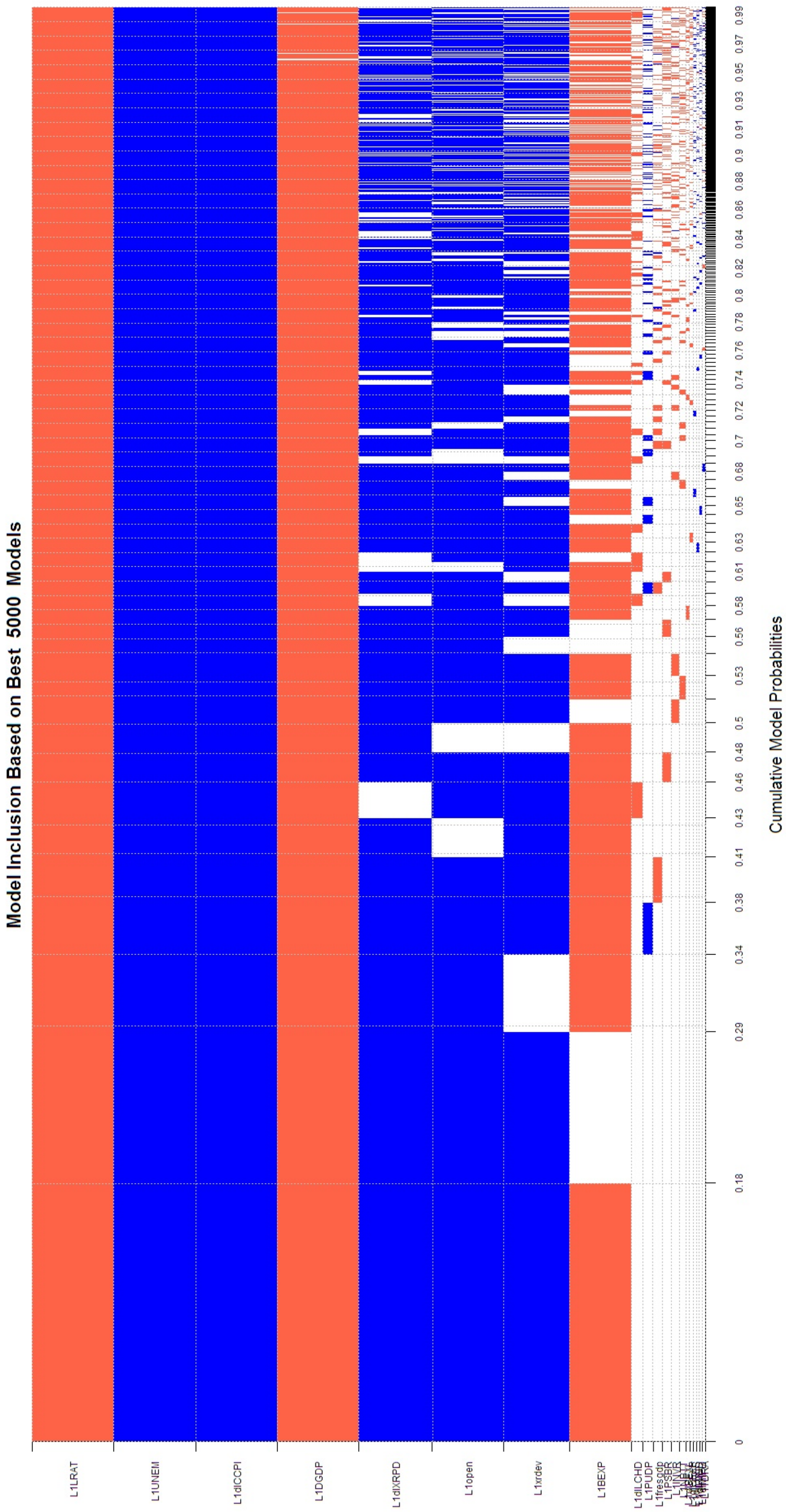


Figure 6.2: g=UIP(BRIC) Model Inclusion

### 6.1.2 g=RIC

This settings uses $g = K^2$ as described in detail in the Subsection 3.10.1, where $K$ is the number of variables or one of the data matrix $X$ dimensions. In our case, $K = 20$, thus $g = 400$.

Again, the *number of models visited* compared to the benchmark setting is low. The *PMP correlation* reaches a high value of 1.0000 which can be seen in the Table 6.3.

The model size tends to be substantially lower than the benchmark one, the *Mean number of regressors* in the Table 6.3 shows 8.5473 which is about the same size as the previous case where $g = UIP(BRIC)$ (Subsection 6.1.1). The uniformely distributed model prior size again turns out to be rather normally distributed and very similar to the previous $g = UIP(BRIC)$ case, as depicted in the Figure 6.3.

Following the Figure 6.4, the model inclusion again amplifies the relative importance of the best spread-describing variables from the benchmark version.

The coefficients are shown in the Table 6.4.

| | | |
|---|---|---|
| Mean no. regressors | Draws | Burnins |
| "8.5473" | "4e+06" | "2e+06" |
| Time | No. models visited | Modelspace 2^K |
| "35.78847 mins" | "917473" | "1e+06" |
| % visited | % Topmodels | Corr PMP |
| "87" | "100" | "1.0000" |
| No. Obs. | Model Prior | g-Prior |
| "616" | "random / 10" | "RIC" |
| Shrinkage-Stats | | |
| "Av=0.9975" | | |

Table 6.3: g=RIC Summary

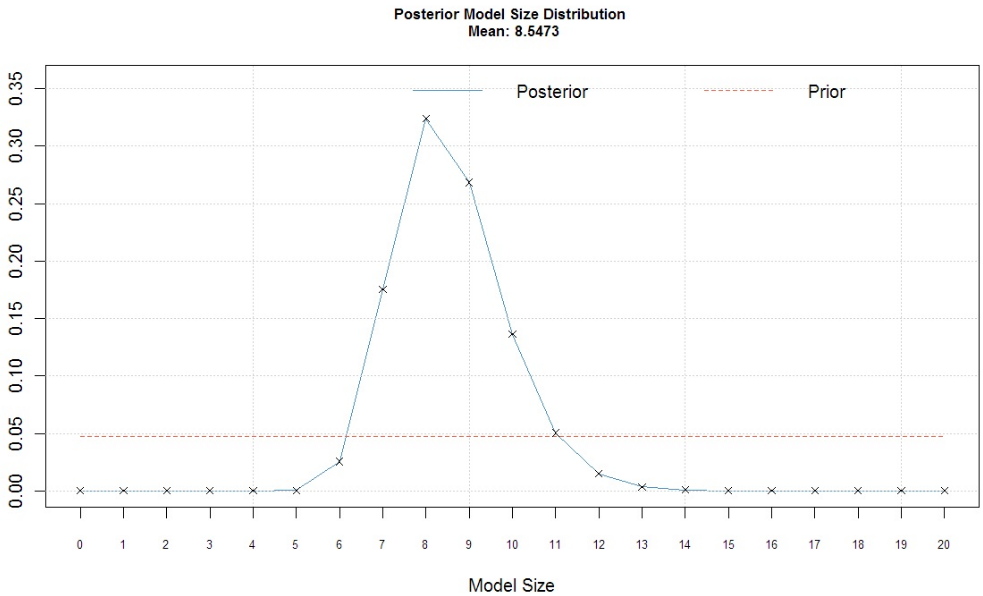


Figure 6.3: g=RIC Model Size

| Variable | Variable Description | PIP | Post Mean | Post SD | Pos. | Idx |
|---|---|---|---|---|---|---|
| L1LRAT | lending interest rate | 1.000 | -0.156 | 0.013 | 0.000 | 4 |
| L1UNEM | unemployment | 1.000 | 0.170 | 0.028 | 1.000 | 9 |
| L1dlCCPI | inflation | 1.000 | 21.621 | 2.981 | 1.000 | 11 |
| L1DGDP | real GDP growth | 0.997 | -0.179 | 0.037 | 0.000 | 2 |
| L1open | openness ratio ($\frac{X+M}{2GDP}$) | 0.902 | 0.944 | 0.484 | 1.000 | 19 |
| L1dlXRPD | exchange rate (domestic/USD) change | 0.897 | 4.375 | 1.822 | 1.000 | 17 |
| L1xrdev | exchange rate deviation from trend | 0.838 | 0.023 | 0.013 | 1.000 | 20 |
| L1BEXP | budget expenditure / GDP | 0.793 | -0.021 | 0.013 | 0.000 | 1 |
| L1PUDP | public debt / GDP | 0.163 | 0.001 | 0.002 | 1.000 | 7 |
| L1fresgdp | foreign exchange reserves (excl.gold) / GDP | 0.154 | -0.198 | 0.616 | 0.070 | 18 |
| L1dlLCHD | labor cost growth | 0.142 | -0.397 | 1.169 | 0.002 | 14 |
| L1PSBR | budget balance / GDP | 0.132 | -0.006 | 0.017 | 0.000 | 6 |
| L1INVR | foreign direct investment / GDP | 0.119 | -0.004 | 0.014 | 0.030 | 3 |
| L1NBTT | terms of trade | 0.098 | -0.001 | 0.002 | 0.000 | 5 |
| L1dlBEXL | budget expenditure growth | 0.061 | -0.081 | 0.474 | 0.000 | 10 |
| L1dlCEXP | exports growth | 0.057 | -0.078 | 0.500 | 0.000 | 12 |
| L1dlFRES | foreign exchange reserves (excl.gold) growth | 0.050 | 0.010 | 0.080 | 1.000 | 13 |
| L1dlRIND | real industry growth | 0.047 | 0.046 | 0.914 | 0.873 | 16 |
| L1dlMIPD | abroad debt service growth | 0.044 | 0.009 | 0.109 | 0.927 | 15 |
| L1TDRA | trade balance / GDP | 0.041 | 0.000 | 0.002 | 0.591 | 8 |

- *Note 1:* All variables were 1 year lagged and are marked by "L1" in their names.
- *Note 2:* "Pos." stands for a conditional probability of the effect sign being positive.

Table 6.4: g=RIC Coefficient Estimates

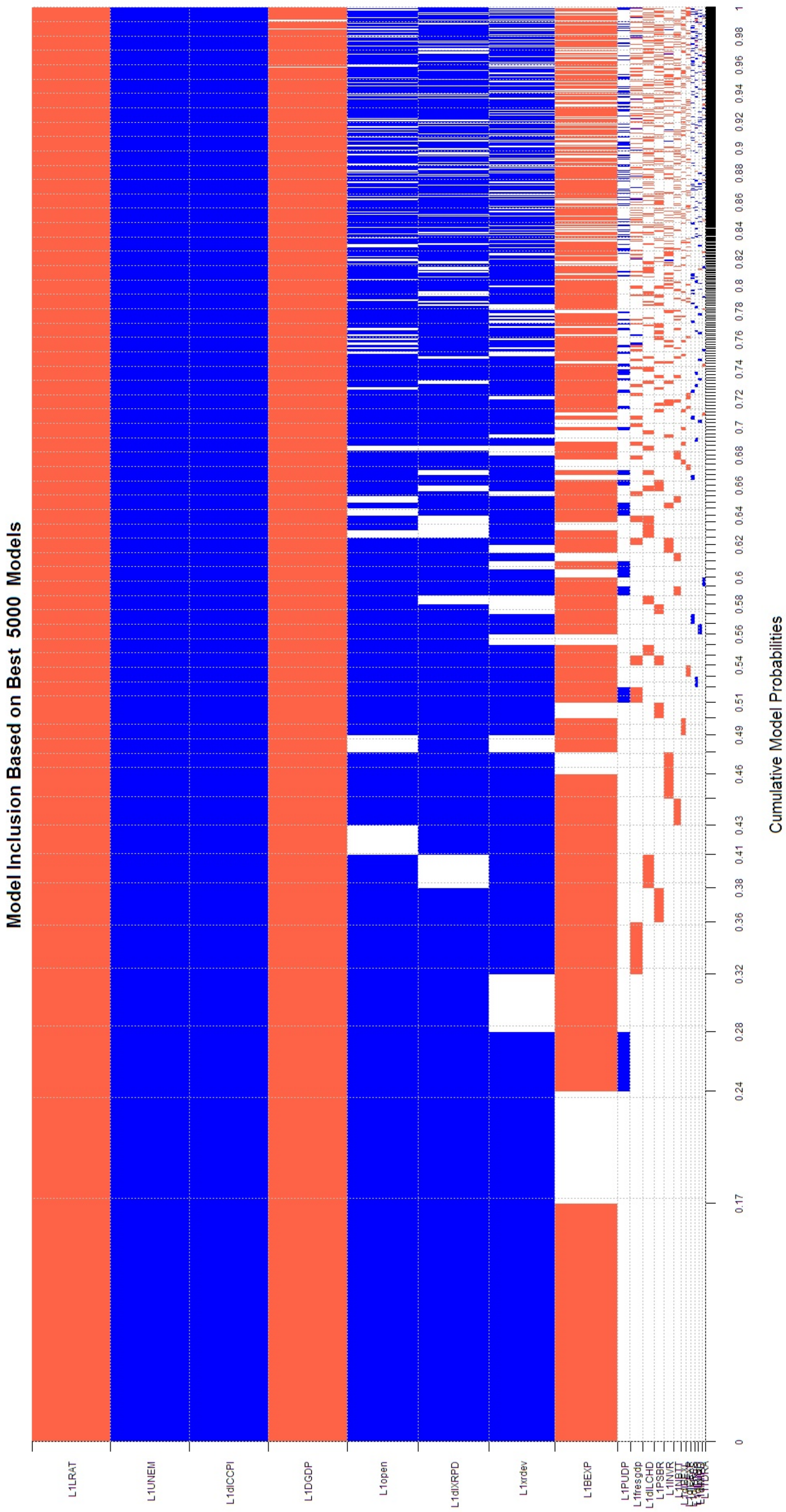


Figure 6.4: g=RIC Model Inclusion

## 6.2 Model size prior sensitivity

This section keeps the setting of $g = hyper - BRIC$ while the different models size priors are applied.

### 6.2.1 model prior = uniform

This settings uses the benchmark $g = hyper - BRIC$ but the *model size prior=random* which is described in the Subsection 3.10.2.

The *number of models visited* is only slightly lower than the benchmark one, the algorithm went through 1953780 models which is lower by only 17.5%. The *PMP correlation* still reaches relatively high value of 0.9997, which can be seen in the Table 6.5.

The model size tends to be close to the benchmark. The *mean number of regressors* in the Table 6.5 shows 10.9695 which is close to the 12.5540. The normally distributed prior model size around the mean exactly in the middle at 10.0000 turns out into posterior model size exerting a slightly positive skew and smaller variance (Figure 6.5).

Following the Figure 6.6, the model inclusion still slightly amplifies the relative importance of the best and medium-good spread-describing variables against worse spread descriptors. The good and medium-good spread descriptors have their PIPs similar with the benchmark.

The coefficients are shown in the Table 6.6.

| | | |
|---|---|---|
| Mean no. regressors | Draws | Burnins |
| "10.9695" | "4e+06" | "2e+06" |
| Time | No. models visited | Modelspace 2^K |
| "55.65415 mins" | "1953780" | "1e+06" |
| % visited | % Topmodels | Corr PMP |
| "186" | "96" | "0.9997" |
| No. Obs. | Model Prior | g-Prior |
| "616" | "uniform / 10" | "hyper (a=2.003247)" |
| Shrinkage-Stats | | |
| "Av=0.9651, Stdev=0.016" | | |

Table 6.5: Uniform Model Prior Summary

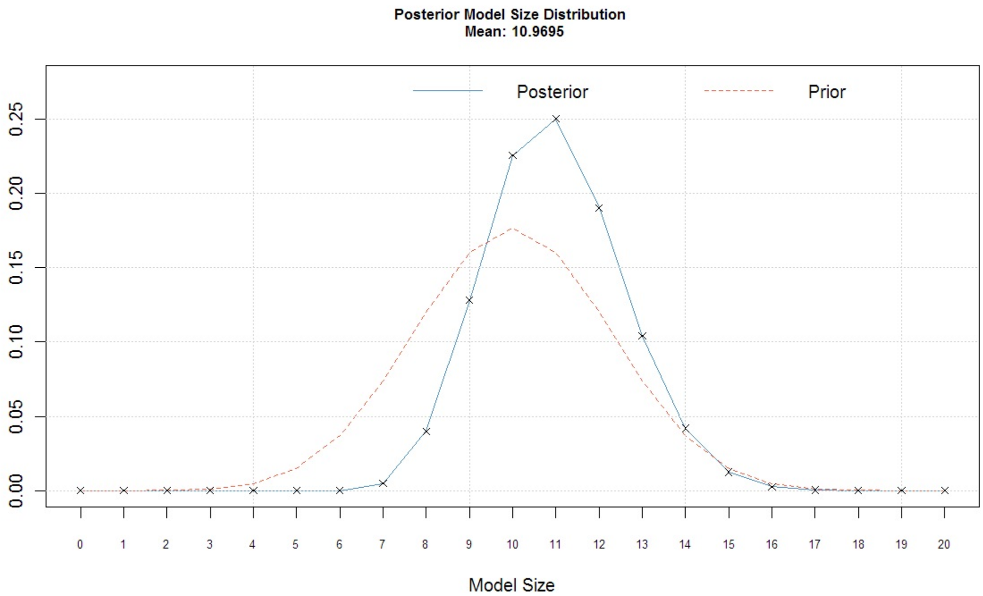


Figure 6.5: Uniform Model Prior Model Size

| Variable | Variable Description | PIP | Post Mean | Post SD | Pos. | Idx |
|---|---|---|---|---|---|---|
| L1LRAT | lending interest rate | 1.000 | -0.150 | 0.012 | 0.000 | 4 |
| L1UNEM | unemployment | 1.000 | 0.162 | 0.028 | 1.000 | 9 |
| L1dlCCPI | inflation | 1.000 | 20.795 | 2.955 | 1.000 | 11 |
| L1DGDP | real GDP growth | 1.000 | -0.173 | 0.040 | 0.000 | 2 |
| L1open | openness ratio ($\frac{X+M}{2GDP}$) | 0.983 | 1.202 | 0.516 | 1.000 | 19 |
| L1xrdev | exchange rate deviation from trend | 0.965 | 0.026 | 0.010 | 1.000 | 20 |
| L1BEXP | budget expenditure / GDP | 0.954 | -0.026 | 0.011 | 0.000 | 1 |
| L1dlXRPD | exchange rate (domestic/USD) change | 0.922 | 4.337 | 1.730 | 1.000 | 17 |
| L1fresgdp | foreign exchange reserves (excl.gold) / GDP | 0.473 | -0.696 | 0.937 | 0.008 | 18 |
| L1PUDP | public debt / GDP | 0.419 | 0.002 | 0.003 | 1.000 | 7 |
| L1PSBR | budget balance / GDP | 0.326 | -0.012 | 0.024 | 0.000 | 6 |
| L1INVR | foreign direct investment / GDP | 0.315 | -0.010 | 0.019 | 0.006 | 3 |
| L1NBTT | terms of trade | 0.267 | -0.001 | 0.003 | 0.000 | 5 |
| L1dlLCHD | labor cost growth | 0.221 | -0.354 | 1.143 | 0.014 | 14 |
| L1dlBEXL | budget expenditure growth | 0.208 | -0.268 | 0.829 | 0.000 | 10 |
| L1dlCEXP | exports growth | 0.201 | -0.268 | 0.880 | 0.000 | 12 |
| L1dlFRES | foreign exchange reserves (excl.gold) growth | 0.192 | 0.045 | 0.159 | 1.000 | 13 |
| L1dlRIND | real industry growth | 0.163 | 0.274 | 1.533 | 0.984 | 16 |
| L1dlMIPD | abroad debt service growth | 0.160 | 0.035 | 0.209 | 0.958 | 15 |
| L1TDRA | trade balance / GDP | 0.146 | 0.000 | 0.004 | 0.550 | 8 |

- *Note 1:* All variables were 1 year lagged and are marked by "L1" in their names.
- *Note 2:* "Pos." stands for a conditional probability of the effect sign being positive.

Table 6.6: Uniform Model Prior Coefficient Estimates

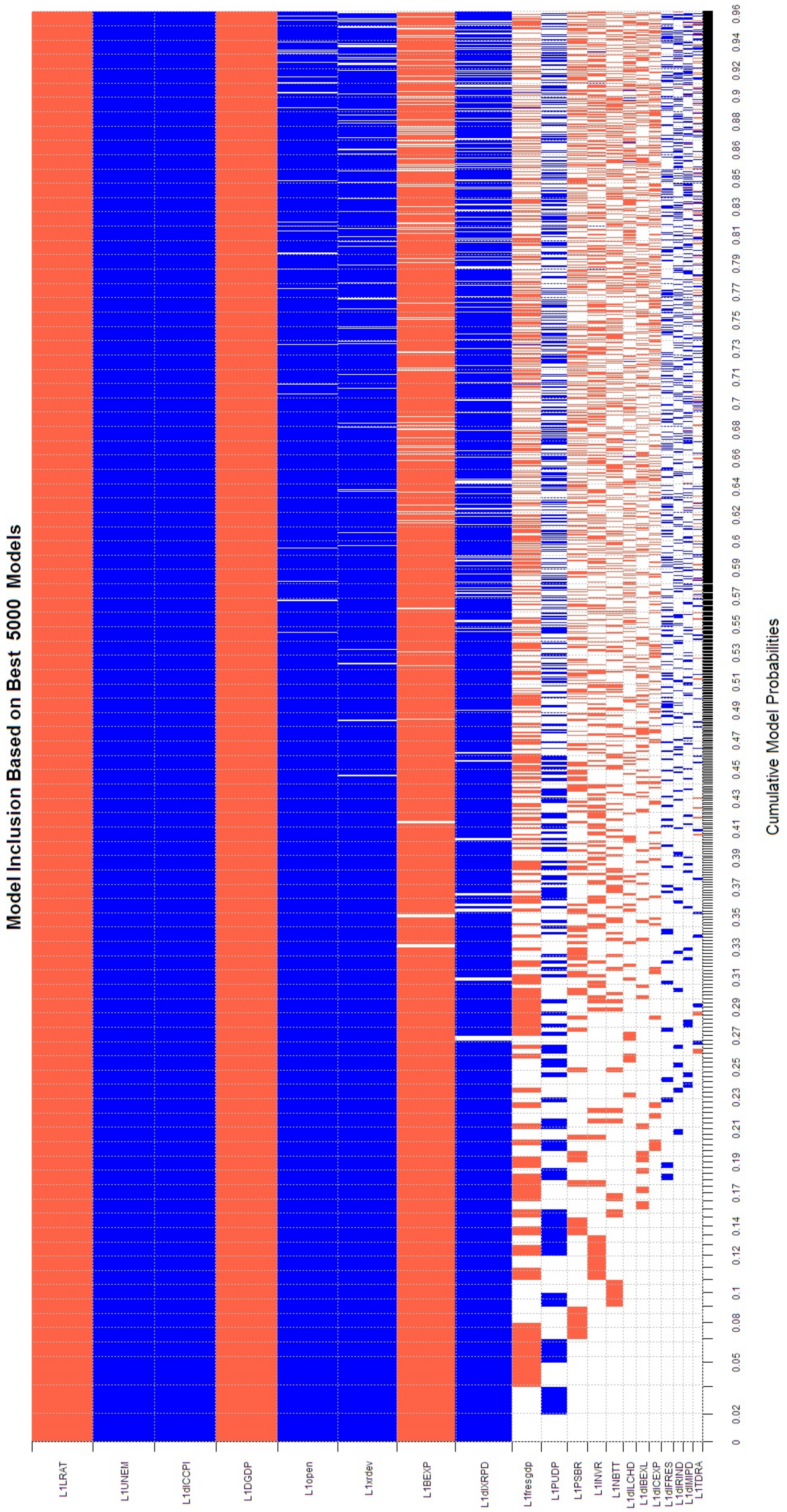


Figure 6.6: Uniform Model Prior Model Inclusion

### 6.2.2 model prior = fixed

The *model size prior*=*fixed* takes place for this section.

The *number of models visited* is only slightly lower than the benchmark one, the algorithm went through 1953773 models which is very the same number as in the Subsection 6.2.1 elliciting *model prior = uniform*. Compared to the benchmark, this number is lower by only 17.5%. The *PMP correlation* still reaches relatively high value and the same as in the Subsection 6.2.1 which is 0.9997 (shown in the Table 6.7).

The model size tends to be close to the benchmark. The *mean number of regressors* in the Table 6.7 shows 10.9707 which is slightly less than the benchmark's 12.5540. As in the previous case (Subsection 6.2.1), the normally distributed prior model size around the mean, exactly in the middle at 10.0000, turns out into posterior model size exerting a slightly positive skew and smaller variance (Figure 6.7).

The Figure 6.8 exerts very similar characteristics as the previous Figure 6.6. the model inclusion still slightly amplifies the relative importance of the best and medium-good spread-describing variables against worse spread descriptors. While the good and medium-good spread descriptors have their PIPs similar with the benchmark.

The coefficients are shown in the Table 6.8.

| | | |
|---|---|---|
| Mean no. regressors | Draws | Burnins |
| "10.9707" | "4e+06" | "2e+06" |
| Time | No. models visited | Modelspace 2^K |
| "57.14062 mins" | "1953773" | "1e+06" |
| % visited | % Topmodels | Corr PMP |
| "186" | "96" | "0.9997" |
| No. Obs. | Model Prior | g-Prior |
| "616" | "fixed / 10" | "hyper (a=2.003247)" |
| Shrinkage-Stats | | |
| "Av=0.9651, Stdev=0.016" | | |

Table 6.7: Fixed Model Prior Summary

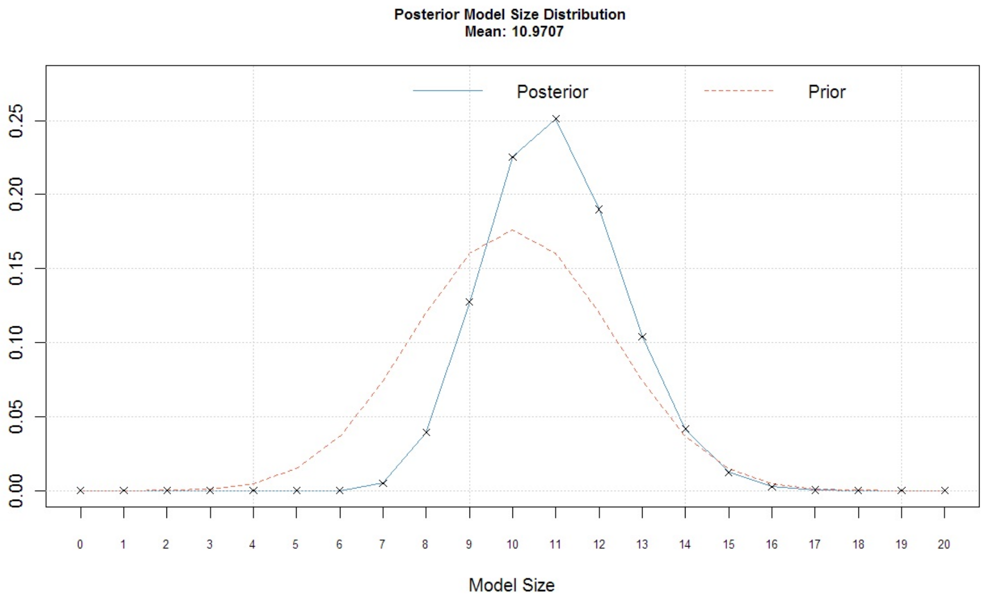


Figure 6.7: Fixed Model Prior Model Size

| Variable | Variable Description | PIP | Post Mean | Post SD | Pos. | Idx |
|---|---|---|---|---|---|---|
| L1LRAT | lending interest rate | 1.000 | -0.150 | 0.012 | 0.000 | 4 |
| L1UNEM | unemployment | 1.000 | 0.162 | 0.028 | 1.000 | 9 |
| L1dlCCPI | inflation | 1.000 | 20.795 | 2.955 | 1.000 | 11 |
| L1DGDP | real GDP growth | 1.000 | -0.173 | 0.040 | 0.000 | 2 |
| L1open | openness ratio $(\frac{X+M}{2GDP})$ | 0.983 | 1.202 | 0.516 | 1.000 | 19 |
| L1xrdev | exchange rate deviation from trend | 0.965 | 0.026 | 0.010 | 1.000 | 20 |
| L1BEXP | budget expenditure / GDP | 0.954 | -0.026 | 0.011 | 0.000 | 1 |
| L1dlXRPD | exchange rate (domestic/USD) change | 0.922 | 4.337 | 1.730 | 1.000 | 17 |
| L1fresgdp | foreign exchange reserves (excl.gold) / GDP | 0.473 | -0.696 | 0.937 | 0.008 | 18 |
| L1PUDP | public debt / GDP | 0.419 | 0.002 | 0.003 | 1.000 | 7 |
| L1PSBR | budget balance / GDP | 0.326 | -0.012 | 0.024 | 0.000 | 6 |
| L1INVR | foreign direct investment / GDP | 0.315 | -0.010 | 0.019 | 0.006 | 3 |
| L1NBTT | terms of trade | 0.267 | -0.001 | 0.003 | 0.000 | 5 |
| L1dlLCHD | labor cost growth | 0.221 | -0.354 | 1.143 | 0.014 | 14 |
| L1dlBEXL | budget expenditure growth | 0.208 | -0.268 | 0.829 | 0.000 | 10 |
| L1dlCEXP | exports growth | 0.201 | -0.268 | 0.880 | 0.000 | 12 |
| L1dlFRES | foreign exchange reserves (excl.gold) growth | 0.192 | 0.045 | 0.159 | 1.000 | 13 |
| L1dlRIND | real industry growth | 0.163 | 0.274 | 1.533 | 0.984 | 16 |
| L1dlMIPD | abroad debt service growth | 0.160 | 0.035 | 0.209 | 0.958 | 15 |
| L1TDRA | trade balance / GDP | 0.146 | 0.000 | 0.004 | 0.550 | 8 |

- *Note 1:* All variables were 1 year lagged and are marked by "L1" in their names.
- *Note 2:* "Pos." stands for a conditional probability of the effect sign being positive.

Table 6.8: Fixed Model Prior Coefficient Estimates

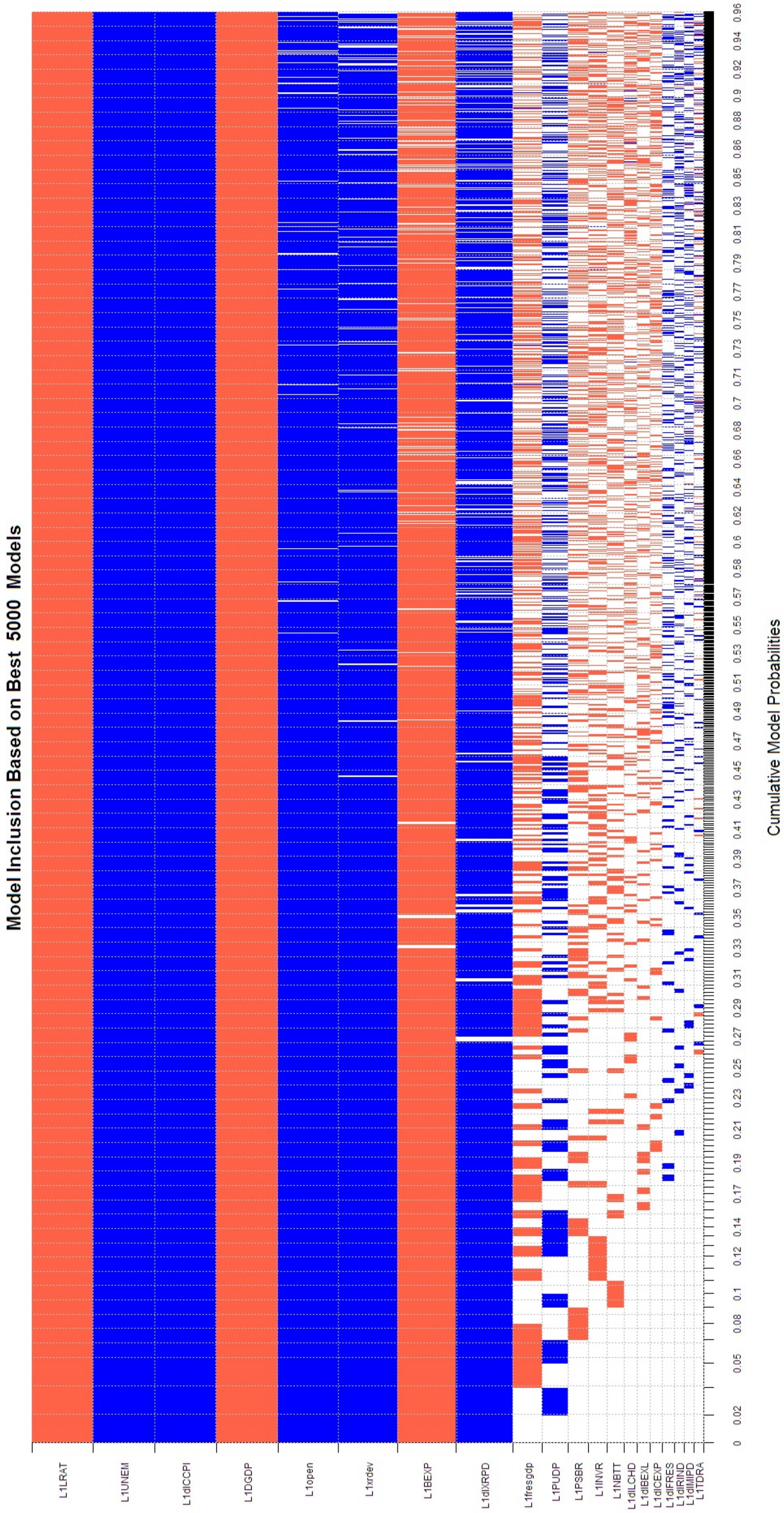


Figure 6.8: Fixed Model Prior Model Inclusion

## 6.3 Sensitivity Analysis

This section aims to sum up the results of the robustness analysis. It uses Figure 6.9 for this purpose. This figure shows the PIPs of the four alternatives as ordered by the benchmark. The fields are colored to ease the graphical representation of the PIPs magnitudes.

The sensitivity analysis shows a high level of the results' consistency over all alternatives. The best spread descriptors remain separated from the medium and low quality spread-explaining variables across all the alternatives. There is a significant gap between the best and medium quality descriptors in the PIP.

The $g - prior$ settings $g = UIP(BRIC)$ and $g = RIC$ undervalues the PIPs of the medium and low quality spread predictors while the best spread descriptors reach the same PIP values as other alternatives. The gap between PIPs for these two alternatives is the most obvious from all the alternatives.

The different priors for the model size cause the PIPs to appear more consistently towards their linear descending trend than the benchmark. Both of the settings, *model prior size*=$\{uniform, fixed\}$ yield very similar PIPs.

The benchmark ellicits relatively the highest PIPs for the medium and low quality spread predictors from all the alternatives. Moreover, these PIPs tend to be more uniform than in the others alternatives.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | **parameter prior g** | hyper-BRIC | UIP(BRIC) | RIC | hyper-BRIC | hyper-BRIC |
| | | **model size prior** | random | random | random | uniform | fixed |
| | | | **BENCHMARK** | | | | |
| **Variable** | **Variable Description** | | **PIP** | **PIP** | **PIP** | **PIP** | **PIP** |
| L1LRAT | lending interest rate | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| L1UNEM | unemployment | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| L1dlCCPI | inflation | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| L1DGDP | real GDP growth | | 1.000 | 0.996 | 0.997 | 1.000 | 1.000 |
| L1open | openness ratio (X+M/2GDP) | | 0.992 | 0.877 | 0.902 | 0.983 | 0.983 |
| L1xrdev | exchange rate deviation from trend | | 0.982 | 0.801 | 0.838 | 0.965 | 0.965 |
| L1BEXP | budget expenditure / GDP | | 0.974 | 0.755 | 0.793 | 0.954 | 0.954 |
| L1dlXRPD | exchange rate (domestic/USD) change | | 0.947 | 0.893 | 0.897 | 0.922 | 0.922 |
| L1fresgdp | foreign exchange reserves (excl.gold) / GDP | | 0.621 | 0.117 | 0.154 | 0.473 | 0.473 |
| L1PUDP | public debt / GDP | | 0.537 | 0.128 | 0.163 | 0.419 | 0.419 |
| L1INVR | foreign direct investment / GDP | | 0.460 | 0.098 | 0.119 | 0.315 | 0.315 |
| L1PSBR | budget balance / GDP | | 0.451 | 0.108 | 0.132 | 0.326 | 0.326 |
| L1NBTT | terms of trade | | 0.391 | 0.079 | 0.098 | 0.267 | 0.267 |
| L1dlCEXP | exports growth | | 0.352 | 0.045 | 0.057 | 0.201 | 0.201 |
| L1dlBEXL | budget expenditure growth | | 0.345 | 0.047 | 0.061 | 0.208 | 0.208 |
| L1dlFRES | foreign exchange reserves (excl.gold) growth | | 0.343 | 0.038 | 0.050 | 0.192 | 0.192 |
| L1dlLCHD | labor cost growth | | 0.323 | 0.137 | 0.142 | 0.221 | 0.221 |
| L1dlMIPD | abroad debt service growth | | 0.294 | 0.033 | 0.044 | 0.160 | 0.160 |
| L1dlRIND | real industry growth | | 0.292 | 0.037 | 0.047 | 0.163 | 0.163 |
| L1TDRA | trade balance / GDP | | 0.270 | 0.032 | 0.041 | 0.146 | 0.146 |

Figure 6.9: PIP Sensitivity Analysis

- *Note 1*: Variables ordered by a descending order with respect to the benchmark model's PIP.
- *Note 2*: The color signifies the PIP size. The fields are in shades of
  - green, for highPIP above 0.75,
  - yellow, for PIP $\in [0.25, 0.75]$,
  - red, for PIP below 0.25.

# Chapter 7

# Conclusion

The purpose of this thesis is twofold. First, to present the BMA method and second, to demonstrate the abilities of this method employing spread explaining variables to predict the spread size.

In this thesis, we use a panel data (Chapter 2) of developed countries (Table 2.1) to study the spread. The set of included variables (Table 2.2) was inspired by the literature as the most frequently used because of its high statistical significance. But as different models in the literature stand for different combinations of variables, the model uncertainty takes place consequently. In this thesis, we try to minimize the model uncertainty using the BMA method.

Compared to standard modelling methods which consist from a chain of several models with decreasing number of included variables based on their statistical significance, the BMA method tries to handle the model uncertainty by evaluation of as many model combinations as possible. These results from single submodels are weighted by the PMP (Section 3.6) and their coefficients are averaged together. Aside of the robust averaging method, an additional interesting property is provided, the PIP, which (Section 3.8) indicates how much is each variable, given the dataset, likely to appear in the true model, as it is shown in the results (Table 5.2).

To relate the results to the literature, we use **positive** sign for spread widening effect which is seen as a bad signal for a country's economy situation while a **negative** sign stands for the spread decrease and is, in normal times, usually seen as a good signal for the economy's future and creditworthiness. We confirm the strong influence in the Table 5.3 as compared with Frankel & Saravelos (2010) and the same direction of the effect on the *spread* which is compared with Rowland & Torres (2004) and Cantor & Packer (1996). The outcome is as follows:

- **positive** influence is identified for *inflation* and *public debt / GDP*,
- **negative** influence on the spread size is found for the *real GDP growth*, *foreign exchange reserves / GDP*.

As other strong spread predictors, we confirm the high number of inclusions made in the overview by Frankel & Saravelos (2010). Namely, we find

- **positive** influence for the *exchange rate trend deviation* and *exchange rate change (domestic/USD)*
- **negative** influence for the *lending interest rate*, *budget expenditure / GDP* and *budget balance / GDP*.

Following the PIP which suggests that several variables' inclusions are rather unlikely, we tend to deny the inclusions of the *terms of trade*, *exports growth*, *budget expenditure growth*, *foreign exchange reserves growth* or *trade balance / GDP* as consistent spread predictors.

In this thesis, we use two variables that despite they were not found statistically significant in the literature these variables outperform most of the others in the BMA framework by their high probability of inclusion (PIP). It is the *openness ratio* ($\frac{X+M}{2GDP}$) and the *unemployment*. Both of these were identified with the same positive and thus spread widening effect.

We expected a country trading the goods in large extent to be a better sign for its competitiveness improving its creditworthiness and consequently decreasing the spread than a non-trading country. The positive —spread widening— effect of the *openness* on the spread that we identified is therefore surprising. The spread, according our results, is supposed to widen as country openness increase. While the spread widening is seen as a bad signal because the country's economy becomes less trustworthy in contrasts with the idea of the international trade, and the exports in particular, being an "international hostage" guaranteeing country's competitiveness.

The sensitivity analysis (Section 6.3) compares the outcomes of the BMA for different priors, both on the parameter and model size spaces. This analysis provides very consistent results across all the alternatives. Moreover, the group of best spread-predicting variables reaches similar PIP values across the alternatives.

The algorithm advanced faster towards high quality estimates in the non-hyper settings for the $g - prior$ (for $g = \{UIP, RIC, BRIC\}$). The number of models visited was only a third compared with the asymptotically consistent $hyper - g$ settings in the benchmark while the quality of the estimation, measured by the PMP corelation was satisfactorily high. The straightforward behavior of the algorithm for non-hyper $g$ settings amplifies the relative importance of the best descriptors at the expense of minimizing the relative importance of worse predictors.

The effect of different *model size priors*, using $\{uniform, fixed\}$ settings caused a 15% decrease of the PIP for both the medium-good and worse spread predictors.

The Figure 5.4 shows selected best models (with the highest PMPs) covering a substantial part of the model probability space. This figure visualizes the mass of variables' inclusions and ellicits this way the spread-describing capabilities of the variables. This is how the BMA method controls for the model uncertainty. Contrasting to standard modelling methods that have no explicit control for it. This figure, together with sensitivity analysis (Section 6.3) is a test of the results' consistency justifying the use of the BMA method in the field of model uncertainty.

# Bibliography


Abiad, A. (2003): "Early Warning Systems: A Survey and a Regime-Switching Approach." *IMF Working Papers 03/32*, International Monetary Fund.

Alessi, L. & C. Detken (2009): "'Real Time' early warning indicators for costly asset price boom/bust cycles - a role for global liquidity." *Working Paper Series 1039*, European Central Bank.

Baldacci, E. (2007): "Beyond the Davos Consensus: A New Approach to Country Risk." *Technical report*, SACE Working Paper, 3.

Baldacci, E., S. Gupta, & A. Mati (2008): "Is it (Still) Mostly Fiscal? Determinants of Sovereign Spreads in Emerging Markets." *IMF Working Papers 08/259*, International Monetary Fund.

Cantor, R. & F. Packer (1996): "Determinants and Impact of Sovereign Credit Ratings." *Economic Policy Review* **(Oct)**: pp. 37–53.

Crespo Cuaresma, J. & T. Slacik (2009): "On the determinants of currency crises: The role of model uncertainty." *Journal of Macroeconomics* **31(4)**: pp. 621–632.

Doppelhoffer, G. & M. Weeks (2008): "Robust Model Averaging." *Technical report*, CESifo Conference, 14-15 March 2008.

Edwards, S. (1984): "LDC's Foreign Borrowing and Default Risk: An Empirical Investigation." *NBER Working Papers 1172*, National Bureau of Economic Research, Inc.

Eichengreen, Barry & Rose, A. K. . W. C. (October 1995): "Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks." *Economic Policy, Blackwell Publishing* **Vol.10, No. 21,**: pp. p.249–312.

EICHENGREEN, B. & A. MODY (1998): "What Explains Changing Spreads on Emerging-Market Debt: Fundamentals or Market Sentiment?" *NBER Working Papers 6408*, National Bureau of Economic Research, Inc.

EICHER, T. S., A. LENKOSKI, & A. RAFTERY (2009): "Bayesian Model Averaging and Endogeneity Under Model Uncertainty: An Application to Development Determinants." *Working Papers UWEC-2009-19*, University of Washington, Department of Economics.

EICHER, T. S., C. PAPAGEORGIOU, & A. E. RAFTERY (2011): "Default priors and predictive performance in Bayesian Model Averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

FELDKIRCHER, M. & S. ZEUGNER (2009): "Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging." *IMF Working Papers 09/202*, International Monetary Fund.

FERNANDEZ, C., E. LEY, & M. F. J. STEEL (2001): "Benchmark priors for Bayesian Model Averaging." *Journal of Econometrics* **100(2)**: pp. 381–427.

FRANKEL, J. A. & A. K. ROSE (1996): "Currency Crashes in Emerging Markets: Empirical Indicators." *CEPR Discussion Papers 1349*, C.E.P.R. Discussion Papers.

FRANKEL, J. A. & G. SARAVELOS (2010): "Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability? Evidence from the 2008-09 Global Crisis." *NBER Working Papers 16047*, National Bureau of Economic Research, Inc.

HAWKINS, J. & M. KLAU (2000): "Measuring potential vulnerabilities in emerging market economies." *BIS Working Papers 91*, Bank for International Settlements.

HILSCHER, J. & Y. NOSBUSCH (2010): "Determinants of Sovereign Risk: Macroeconomic Fundamentals and the Pricing of Sovereign Debt." *Review of Finance* **14(2)**: pp. 235–262.

KAMIN, S. B. & K. VON KLEIST (1999): "The evolution and determinants of emerging market credit spreads in the 1990s." *International Finance Discussion Papers 653*, Board of Governors of the Federal Reserve System (U.S.).

KAMINSKY, G., S. LIZONDO, & C. M. REINHART (1998): "Leading Indicators of Currency Crises." *IMF Staff Papers* **45(1)**: pp. 1–48.

KOOP, G. (2003): *Bayesian Econometrics*. John Wiley.

MCGUIRE, P. & M. A. SCHRIJVERS (2003): "Common factors in emerging market spreads." *BIS Quarterly Review* .

MIN, H. G. (1998): "Determinants of emerging market bond spread : do economic fundamentals matter?" *Policy Research Working Paper Series 1899*, The World Bank.

MORAL-BENITO, E. (2010a): "Model Averaging In Economics." *Working Papers wp2010-1008*, CEMFI.

MORAL-BENITO, E. (2010b): "Panel Growth Regressions With General Predetermined Variables: Likelihood-Based Estimation And Bayesian Averaging." *Working Papers wp2010 1006*, CEMFI.

OBSTFELD, M., J. C. SHAMBAUGH, & A. M. TAYLOR (2009): "Financial Instability, Reserves, and Central Bank Swap Lines in the Panic of 2008." *American Economic Review* **99(2)**: pp. 480–86.

della PAOLERA, G. & M. GRANDES (2007): "The True Measure of Country Risk: A Primer on the Interrelations between Solvency and the Polity Structure of Emerging Markets, Argentina 1886-1892." In "The Decline of Latin American Economies: Growth, Institutions, and Crises," NBER Chapters, pp. 195–212. National Bureau of Economic Research, Inc.

ROSE, A. K. & M. M. SPIEGEL (2009): "Cross-country causes and consequences of the 2008 crisis: early warning." *Working Paper Series 2009-17*, Federal Reserve Bank of San Francisco.

ROSE, A. K. & M. M. SPIEGEL (2011): "Cross-country causes and consequences of the crisis: An update." *European Economic Review* **55(3)**: pp. 309–324.

ROWLAND, P. (2004): "Determinants of Spread, Credit Ratings and Creditworthiness for Emerging Market Sovereign Debt: A Follow-Up Study Using Pooled Data Analysis." *Borradores de Economia 296*, Banco de la Republica de Colombia.

Rowland, P. & J. L. Torres (2004): "Determinants of Spread and Creditworthiness for Emerging Market Sovereign Debt:A Panel Data Study." *Borradores de Economia 295*, Banco de la Republica de Colombia.

Uribe, M. & V. Z. Yue (2006): "Country spreads and emerging countries: Who drives whom?" *Journal of International Economics* **69(1)**: pp. 6–36.

Zellner, A. (1986): "A tale of forecasting 1001 series : The Bayesian knight strikes again." *International Journal of Forecasting* **2(4)**: pp. 491–494.

Zeugner, S. (2011): "Bayesian Model Averaging with BMS." *Technical report*, `http://bms.zeugner.eu/` May5, 2011.

# Appendix A

# Content of Enclosed Disc

There is a disc enclosed to the printed copy of this thesis and/or a zipped file attached in the university information system which contain empirical data and source codes.

- folder **excel**: Excel datasheet
- folder **r**: R data code and R workspace
- folder **stata**: Stata data and Stata code
- folder **tex**: LaTeX source code
- thesis PDF version

# Master Thesis Proposal

| | |
|---|---|
| **Author** | Bc. Vojtěch Seman |
| **Supervisor** | PhDr. Marek Rusnák |
| **Proposed topic** | Spread Determinants and Model Uncertainty: A Bayesian Model Averaging Analysis |

**Topic characteristics** The problem of bubbles - their existence and identification is on. The problem it faces is that the people sometimes tend to be unpredictable and goes panic or herding. However, these are only the reactions on informations. In some extent, these informations are based on true data, facts and stories. In this paper we want to relate hard macroeconomic data with the sovereign bond rate as a representation of the relation between an investor and a country as a production unit. The motivation is to assess more robust estimates as an extension to existing literature by defining a new variable, not commonly used in the early warning system literature. In this paper we would like to establish a link between two types of literature. While the first type, known as early warning systems serves as a baseline, with the second, political economy type. Especially, the latter stresses on the importance of a country competitiveness in the long term by the share of exports allocated in sectors that are hardly transferable. This element, in terms of international trade, means a lower probability to lose a comparative advantage which may serve as a safer collateral for the investors. Thus, we will test the significance of disaggregated export groups based on the literature specifications.

In part two, we will try to assess an ideal state of country openness compared to its economic size.

Third part will be dedicated to the determination of the model explaining the sovereign debt to GDP by country macroeconomic properties. As this topic of Economics generates answers using numerical methods and there are no answers without questions, I would like to focus in my thesis on the side generating the questions too. And add an alternative topic as a possible extension. However these two topics are linked together through the world of investment, credit and country competitiveness. This topic matches my specialization and even though there will be probably no numerical models used, this will fit the faculty requirements on the thesis size.

In the next topic, I will react to signals about failures of the project financed by the European Investment Bank. I will provide an institutional description of the quality assessment of the investment made by the EIB with respect to its political objectives.

**Hypotheses** Hypothesis #1: Can country exports serve as an imaginary collateral on the financial markets, in other words: does the higher trade openness prevent a country from going bankrupt? More precisely, exerce disaggregated export groups statistical significance and importance while assessing the sovereign bond rate.

Hypothesis #2: Is there an ideal combination of country properties in the international trade competition? What should a country look like? Countries trade because their productions are different. Would it be reasonable to follow a universal pattern for all of them?

Hypothesis #3: Is the sustainable level of sovereign debt to GDP predictable by the country properties ex ante? Are we able to determine it?

**Methodology** Bayesian model averaging

Binary and multinomial probit and logit regressions

**Outline** Introduction

Literature survey

Sovereign bond spread as the probability of a country going bankrupt - a model

Ideal state of country competitiveness and ideal size of the economic fundamentals

Appropriate level of sovereign debt to GDP relatively to country fundamentals

**Summary**

1. Introduction
2. Theoretical Background

3. Related Work
4. The Model
5. Empirical Verification
6. Conclusion

In the first part of the thesis, I will focus on the macroeconomic description of OFDI, their development and regional distribution in particular. Such a view is useful to see the phenomenon of OFDI in a perspective of the development of the economy as a whole and to be able to compare it potentially with other economies.

## Core bibliography


1. Rose, A. K. & M. M. Spiegel (2009): Cross-country causes and consequences of the 2008 crisis: early warning. *Working Paper Series 2009-17*, Federal Reserve Bank of San Francisco.
2. Hilscher, J. & Y. Nosbusch (2010): Determinants of sovereign risk: Macroeconomic fundamentals and the pricing of sovereign debt. *Review of Finance* **14(2)**: pp. 235–262.
3. della Paolera, G. & M. Grandes (2007): The true measure of country risk: A primer on the interrelations between solvency and the polity structure of emerging markets, argentina 1886-1892. In The Decline of Latin American Economies: Growth, Institutions, and Crises, NBER Chapters, pp. 195–212. National Bureau of Economic Research, Inc.
4. Mody, A. & A. Carare (2010): Spillovers of domestic shocks: Will they counteract the "great moderation"? *IMF Working Papers 10/78*, International Monetary Fund.
5. Biggs, M., T. Mayer, & A. Pick (14 May 2010): The myth of the Phoenix miracle. *Technical report*, `http://www.voxeu.org/index.php?q=node/5038`.
6. De Grauwe, P. (19 May 2010): Fighting the wrong enemy. *Technical report*, `http://www.voxeu.org/index.php?q=node/5062`.


Author

Supervisor