Truths and Myths about the RMB Misalignment: A Meta-Analysis

Yin-Wong Cheung
Shi He
This Version: November 2018 (Please Do Not Cite)

Abstract
We conduct a meta-regression analysis of 69 studies that generated 937 RMB misalignment estimates. The Bayesian Model Averaging approach is adopted to allow for model selection and sampling uncertainties in assessing effects of study characteristics on these RMB misalignment estimates. It is found that the misalignment estimate can be influenced by eight selected study characteristic types; some display positive effects and some negative effects. The RMB misalignment estimate from models with various hypothetical combinations of study characteristics, however, is mostly insignificantly different from zero. It is also shown that the set of significant study characteristics is sensitive to the use of the least squares estimation method, and the choice of benchmark study characteristics.

JEL Codes: C83; F31; F41
Keywords: Bayesian Model Averaging; Clustering Effects; Median Probability Model; RMB Undervaluation; Study Characteristics

Acknowledgments: We would like to thank Menzie Chinn, JF Carpantier, David-Jan Jansen, Juilus Loermann, and participants at the 12th MIFN International Workshop for their comments and suggestions. We also thank Iikka Korhonen and Shaofang Yu for sharing their data. Cheung and He gratefully acknowledge the Hung Hing Ying and Leung Hau Ling Charitable Foundation (孔慶熒及梁巧玲慈善基金) for their support through the Hung Hing Ying Chair Professorship of International Economics (孔慶熒講座教授(國際經濟)).

Correspondence Addresses:
Yin-Wong Cheung, Hung Hing Ying Chair Professor of International Economics, Department of Economics and Finance, City University of Hong Kong, HONG KONG. E-mail: yicheung@cityu.edu.hk.
Shi He, School of Economics, Huazhong University of Science and Technology, Wuhan, China, and Department of Economics and Finance, City University of Hong Kong, HONG KONG. E-mail: heshi818@hotmail.com

1. Introduction

Currency misalignment is a recurring controversial and sentimental topic in the global economy. A recent example is the dispute over the valuation of the Chinese currency, the renminbi (RMB). Since the world witnessed the rapid growth of Chinese trade surplus in the early 2000s, both academics and policymakers have scrutinized China's foreign exchange policy. A typical viewpoint is that the undervalued RMB is the culprit of China's ballooning surplus, and the source of global current account imbalances. The phenomenon has triggered a wave of studies estimating the degree of the RMB misalignment.

The reported empirical estimates of RMB misalignment, after an initial phase dominated by undervaluation estimates, have spanned a rather wide range of over- and undervaluation estimates. The wide dispersion of misalignment estimates is, not surprisingly, similar to different empirical estimates reported for other economic issues.

There are reasons for these misalignment estimates to vary with, say, the choice of models and sample periods. For instance, Cheung *et al.* (2007) points out the potential difficulty of pinning down the magnitude of misalignment due to the absence of a consensual equilibrium exchange rate model.[1] Dunaway *et al.* (2009), on the other hand, reports that the RMB misalignment estimate is quite sensitive to the assumptions underlying estimation methods. Based on the then available empirical studies, Bineau (2010) and Korhonen and Ritola (2011) found that the reported RMB misalignment estimates are associated with some characteristics of these studies including the definition of RMB exchange rate, the choice of theoretical frameworks and estimation methods, publication types, and author's affiliations.

The theme of the current study is not on the choice of the equilibrium exchange rate model and, hence, the appropriate empirical measure of currency misalignment. Instead, we focus on the variation of RMB misalignment estimates in terms of the association between their heterogeneity and characteristics of the corresponding studies. In a sense, we extend the meta-analyses conducted by Bineau (2010) and Korhonen and Ritola (2011), and investigate whether the empirical RMB misalignment estimates vary systematically with one or more study characteristics.

The current study extends the existing literature along several dimensions. First, to the extent possible, we include in our meta-analysis exercise studies that report RMB misalignment

[1] The seminal article Meese and Rogoff (1983) documented the inability of economic models to explain exchange rate movements; a finding that is echoed by, for example, Cheung *et al*. (2005, 2018) and Rossi (2013).

estimates and their corresponding study characteristics. Our sample includes 69 studies that give 937 RMB misalignment estimates. The wide coverage of studies shall minimize selection biases.[2] Further we consider 13 study characteristics that may explain the study-to-study heterogeneity of misalignment estimates. While comparable to those considered by, for example, Bineau (2010), Égert and Halpern (2006), and Korhonen and Ritola (2011), these study characteristics offer a rather extensive coverage.

Second, the Bayesian Model Averaging (BMA) approach is employed to investigate the impact of selected study characteristic types. While the meta-analysis is designed to explore the links between the variable of interest (RMB misalignment estimates in the current case) and the attributes of empirical studies, there is no strong theory on what the key study characteristic types should be. In practice, there is considerable latitude for determining the set of explanatory factors. In choosing a specification from possible combinations of study characteristics, we have to account for, in addition to the usual estimation/sampling uncertainty, the model selection uncertainty. The BMA approach is an established procedure that explicitly accounts for the uncertainties due to both model selection and estimation procedure in generating inference on the parameter estimates of interest.

In anticipation of results, our BMA-based meta-analysis shows that the reported RMB misalignment estimates are affected by some study characteristics. For instance, the study characteristic types include the use of PPP-based data, a cross-sectional setting, and a structural setup tend to yield a strong evidence of RMB undervaluation, while the use of data on real effective exchange rate or nominal (effective) exchange rate, studies prepared by a group of authors from academics, government, and industry, and the estimates for the period of 2009 to 2014 tend to be associated with a weak evidence of RMB undervaluation. However, the use of the least squares regression technique; with and without controlling for clustering on study, can yield different sets of significant study characteristic types. The significance of study characteristics is also sensitive to the choice of benchmark study characteristics.

The remainder of the paper is organized as follows. Section 2 presents the sample of studies, the study characteristics used to explain the variability of study-to-study RMB misalignment estimate, and some graphical analyses. Section 3 reports the results from the BMA-

[2] Bineau (2010) and Korhonen and Ritola (2011) cover, respectively, 17 and 30 studies, and 130 and 99 misalignment estimates in their exercises.

based meta-analysis. Section 4 provides additional results that are obtained from adopting the least squares approach with and without controlling for clustering on study, and from an alternative choice of study characteristics. Section 5 offers some concluding remarks.

2. Preliminaries

Since China launched its reform initiative in 1978, it has experienced an amazingly strong growth performance and rapidly integrated with the global economy; these phenomenal progresses have substantial implications for both the Chinese economy and the world. China's quick ascendance on the global stage has put its RMB policy in the limelight, and has triggered considerable interest in estimating the equilibrium value and the misalignment of the Chinese currency. Chou and Shih (1998) presents an early empirical study on RMB misalignment. The debate on the RMB valuation gathered momentum in 2005 when China's trade surplus surpassed 4% of its GDP and continued to grow strongly to above 8% in 2008. Study on RMB misalignment has slow down a bit in the second half of the 2010s – Almas *et al.* (2017), Cheung *et al.* (2017), and Giannellis and Koukouritakis (2018) are three recent journal articles on RMB misalignment.

In the last two decades, researchers produced a plethora of empirical studies on RMB misalignment that were based on different models and estimation methods, and covered different time periods. These empirical studies have generated conflicting inferences about the level of RMB misalignment; the misalignment estimates range from substantial undervaluation to large overvaluation. We adopt the meta-analysis approach to study the implications of study characteristics for the observed heterogeneity of study-to-study RMB misalignment estimates.

2.1 The Sample of Studies

To alleviate selection bias, we to the extent possible include in our sample studies that report quantitative inferences about the RMB misalignment. The raw sample of studies was constructed via the following means. For studies in English, we include studies considered by Bineau (2010) and Korhonen and Ritola (2011). Then, we conducted searches on the Google Scholar website using the keywords "RMB misalignment," "RMB equilibrium exchange rate," "RMB undervaluation," "RMB overvaluation," "the Chinese currency misalignment," "estimating the Chinese currency," and "RMB valuation." For this set of studies, we identified five early and well-cited studies and collected papers that cited any one of these five studies.[3] For studies in Chinese, we sorted through the top 10 Chinese journals on economics.[4] Last, we went through the studies published between 2015 and 2018 obtained from the previous steps and looked for relevant references.[5] In total, we have 283 studies in the preliminary sample.

From the preliminary sample, we looked for studies that report quantitative estimates of RMB misalignment on or after 1994, and identified a total of 69 studies of which 62 are in English publications and 7 Chinese publications. These studies constituted the sample of studies examined in the following analyses, and are listed in the Appendix A.1.

Two remarks are in order. First, we label the sample of quantitative estimates of RMB misalignment $\mathbf{Y} = \{ Y_i \}$. Specifically, the **Y**-sample has 937 observations on percentage misalignment estimates given by the difference of the actual and the (estimated) equilibrium exchange rate in percentage.[6] In the following, the terms "RMB misalignment estimates" or "misalignment estimates" are used to refer to these quantitative estimates of RMB misalignment.

Second, we focus on misalignment estimates on or after 1994 because China in January 1994 instituted a major change of its exchange rate policy – it replaced a dual exchange rate arrangement with one that managed its currency value against the US dollar. In our sample, 2014 is the last year that we have RMB misalignment estimates.[7]

Figure 1 displays the frequencies of the studies according to the years that they were published. There are a few studies before the 21st century. Chou and Shih (1998) and Zhang (2000) are the two journal articles published between 1998 and 2000. A large proportion of the selected studies published between 2007 and 2013 – the period after China's trade surplus reached high levels. The latest studies in our sample are Almas *et al.* (2017), Cheung *et al.* (2017), and Giannellis and Koukouritakis (2018).

---

3 The identified studies are Chang and Qin (2004), Cheung *et al*. (2007), Chou and Shih (1998), Funke and Rahn (2005), and Zhang (2001); and these studies have garnered, respectively, 122, 288, 158, 224, and 224 citations (Google Scholar, assessed as of June 2018).

4 The 10 Chinese journals are "中国社会科学," "管理世界," "经济研究," "经济学季刊," "世界经济," "金融研究," "中国工业经济," "数量经济学与技术经济研究," "中国农村经济," and "经济学动态," which are considered influential Chinese academic journals on economics.

5 These articles are Agya and Du (2015), Almas *et al.* (2017), Cheung *et al.* (2017), Giannellis and Koukouritakis (2018), Li (2015), and Wang (2015).

6 Misalignment measures calculated from exchange rate data themselves are usually larger than those from exchange rates in logs. Most studies used logged data.

7 Bineau (2010) and Korhonen and Ritola (2011) include misalignment estimates in the periods 1975 to 2008 and 1998 to 2009.

The box plots of RMB misalignment estimates in individual years are presented in Figure 2. For easy reference, the legend of the box plot (which is also known as box and whisker diagram) is given below the figure. Note that a positive misalignment value indicates RMB undervaluation. Figure 2 shows that the median of misalignment estimates displays a slight downward trend between 1994 and 1999, then it trends upward and reaches the largest value in 2004 before drifting lower, and, after a brief pick up in 2010, continues its declining trend toward the end of the sample period. With the exceptions of 2011 and 2014, the majority of misalignment estimates leans toward RMB undervaluation. The within-year variability of these percentage misalignment estimates displays wide and time-varying variations – the year 2004 has the largest inter-quantile range estimate while the year 2001 has the largest range between extreme undervaluation and overvaluation. Further, with the exception of 2009, 2010, 2013, and 2014, there are extreme misalignment estimates reported for individual years in the sample – the occurrence of extreme undervaluation estimates outnumbers extreme overvaluation estimates. In general, the range of year-to-year variability has increased in the early part of the sample, and then reduced over time.[8]

2.2 Study Characteristics

The meta-analysis approach is adopted to investigate the potential roles of features of studies that may explain the variation of RMB misalignment estimates. To this end, we collected information on 13 study characteristics and grouped them into four categories; namely, a) data characteristics, b) theoretical and estimation specifications, c) publication attributes, and d) subsample periods. The definitions of these study characteristics and their corresponding characteristic types are listed in the Appendix A.3.[9]

2.2.1 Data characteristics

The category of data characteristics comprises three study characteristics. Specifically, we coded the studies according to i) whether annual, quarterly, or monthly data are used; ii) whether data are mainly derived from PPP-based data (such as International Comparison Program (ICP) surveys) or not; and iii) whether the RMB real effective exchange rate, RMB real exchange rate

[8] Appendix A.2 presents, for each year, the mean, standard error, minimum, and maximum of misalignment estimates.

[9] For each given study characteristic (e.g. data frequency), its alternative elements/specifications (e.g. "Annual," "Quarterly," and "Monthly") are referred as study characteristic types in this exercise.

against the US dollar, RMB real exchange rate against the Japanese yen or the euro, and other types of exchange rate including nominal (effective) exchange rates are used.

Figure 3 presents, for each data characteristic, the box plots of misalignment estimates across the corresponding study characteristic types. Annual misalignment estimates account for slightly over one half (528/937) of the estimates in the **Y-**sample, and display a high level of variability as indicated by the inter-quantile estimates, the difference between the upper and low adjacent values, and the extreme values (Figure 3.a). The quarterly estimates have the smallest medium and shortest inter-quantile range among these three data frequencies.

Comparing with studies using data derived from market exchange rates, studies using PPP-based data derived from ICP surveys tend to yield a more variable RMB misalignment estimate, and a stronger evidence of undervaluation (Figure 3.b).

Misalignment estimates of the RMB-US dollar real exchange rate and of the RMB real effective exchange rate account for, respectively, 67% and 26% of observations in the **Y-**sample. The RMB-US dollar real exchange rate misalignment estimates have the largest medium value, highest level of volatility, and the most occurrences of extreme (undervaluation) values.

2.2.2 Theoretical and Estimation Specifications

Under the category of theoretical and estimation specifications, there are four study characteristics. For the study characteristic of theoretical settings, we consider five types; namely, “BEERs,” “FEERs,” “Penn effect,” absolute or relative PPP framework (“PPP”), and “Other frameworks.”[10] And, under the three estimation specification characteristics, we classify the studies according to: i) whether a panel, cross-sectional, or time series approach is adopted, ii) whether cointegration framework is used or not, and iii) whether a reduced-form or structural setup is used. The box plots of misalignment estimates of each of these four theoretical and estimation study characteristics are presented in Figure 4.

Over one-third of the misalignment estimates are generated from the BEERs framework (Figure 4.a). The Penn effect regression is one that generates the next most misalignment estimates in the **Y-**sample. These Penn-effect-based estimates, compared with those from other methods,

[10] The other theoretical frameworks include general equilibrium model, recovery mechanism of equilibrium exchange rate, shadow price of foreign exchange approach, and exchange market pressure approach.

yield the largest median of undervaluation estimates, and the largest range of estimates as evidence by the extreme values and the difference between the upper and lower adjacent values.

The median of the estimates from the panel data setting is the largest, followed by the one from cross-sectional setting and then the time series data (Figure 4.b). While studies using time series data general both overvaluation and undervaluation extreme estimates, those using cross-sectional data yield mostly undervaluation extreme estimates.

Figure 4.c indicates that, compared with non-cointegration methods, the use of cointegration approach yields a smaller median and a less volatile estimate of misalignment. The reduced-form approach, compared with a structural approach, generates misalignment estimates that have a smaller median and are more volatile (Figure 4.d).

2.2.3 The publication Attributes

One study characteristic under the category of publication attributes is the venue of publication; that is whether the study is published as an academic journal article, a book chapter, or other forms including as a working paper. Another one is whether the study is published in English or in Chinese.

Figures 5.a and 5.b show that most of the misalignment estimates in our sample are from academic journal articles and from publications in English. Both characteristics are associated with misalignment estimates that display a wide range and numerous extreme values. Under the publication venue study characteristic, book chapters contribute only nine misalignment estimates; however, they give the largest median estimate in this study characteristic.

Another three study characteristics under the category of publication attributes are related to the author(s) of selected studies. One study characteristic is whether any one of the authors has an affiliation with an institution in mainland China. Another study characteristic is whether any one of the authors has a Chinese name (and educated in China at any education level). Lastly, we identify whether all authors in one study have only academic affiliations, only "Government" affiliations,[11] only industry affiliations (e.g. investment bank and commercial bank), or a mix of these affiliations.

[11] "Government" refers to government affiliations (e.g. central banks) or think tanks (e.g. Peterson Institute for International Economics) or international organizations (e.g. IMF and Asian Development Bank).

Figures 5.c and 5.d show that studies that authors who are not affiliated with a mainland China institution or non-Chinese tend to report a more severe degree of RMB undervaluation as indicated by median estimates. On the other hand, a relative large proportion of extreme misalignment estimates are found among studies that have authors who are affiliated with a mainland China institution or Chinese.

Among the four types of author's affiliations, the academic type accounts for slightly over one-half of the percentage misalignment estimates considered here (Figure 5.e). The misalignment estimates presented by studies that were authored by only academics include quite a number of extreme observations though the median is quite small relative to those from other affiliation types.

2.2.4 Subsample periods

China modified its exchange rate policy a few time during our sample period. For instance, it replaced a dual exchange rate arrangement with a policy targeting the US dollar in January 1994, adopted a *de facto* dollar-peg arrangement after the 1997 Asian Financial crisis, migrated to a "managed floating exchange rate regime" in July 2005, reverted back to a stable RMB/dollar rate policy in the midst of the Global financial crisis in 2008, returned to the "managed floating exchange rate regime" in 2011. A pegged rate policy arguably hinders the exchange rate adjustment process and, thus, have implications for currency misalignment. To assess exchange rate policy effects, we investigate if different levels of RMB misalignment estimates are observed in subsample periods 1994-1997, 1998-2004, 2005-2008, 2009-2010, and 2011-2014.

The box plots in Figure 6 show that the periods 1998-2004 and 2005-2008 garner the two highest proportions of observations, and exhibit considerably variable misalignment estimates. The median of the 2011-2014 period is visually smaller than medians from other subsample periods. Indeed, the ratio of undervaluation to non-undervaluation estimates is one for the period 2011-2014, and is the smallest amongst these five subsample periods.

The box plots offer some circumstantial evidence on implications of study characteristics for the RMB misalignment estimates. The observed differential effects across characteristic types of a given study characteristic, however, can be influenced by the interaction of the effects of all study characteristics rather than a single characteristic on misalignment estimates. In the next section, we present some vigorous statistical analyses on the effects of study characteristics.

3. Data Analyses

The figures presented in the previous section are suggestive of the RMB misalignment estimates are associated with some study characteristics. An astute reader will require some additional statistical evidence to confirm the roles of these study characteristics. In the following we adopt the regression framework

$$Y_i = \alpha + \sum_{j=1}^{J} \beta_j X_{ij} + \varepsilon_i \qquad (1)$$

to examine the study-to-study variation of RMB misalignment estimates. In (1) the dependent variable Y_i is the *i*-th RMB misalignment estimate in percentage, the explanatory variable X_{ij} is the *j*-th characteristic type of the study that reported Y_i, and J is the total number of study characteristic types under consideration.

These explanatory variables X_{ij}'s, which are also known as moderator variables in meta-analysis, are qualitative variables that take the form of a zero-one dummy variable. For a given study characteristic say, data frequency, the inclusion of all three data frequency types (i.e., "Annual," "Quarterly," and "Monthly") leads to perfect collinearity in the presence of a regression intercept term (or other qualitative response variables). Because of this, we define for a given study characteristic a "benchmark" characteristic type, which is used as a reference point for assessing the effects of study characteristics. In this Section, for each study characteristic, we identify its benchmark type as the study characteristic type that is adopted by most studies. The 13 selected benchmark characteristic types are identified by the bold-font in Appendix A.3.[12]

Not counting the 13 benchmark types, there are 26 possible explanatory variables in our exercise. In principle there are 2^{26} (= 67,108,864**)** possible empirical specifications. Which set of variables should be included in the empirical analysis? Arguably, despite some anecdotal evidence from the figures presented in the previous section, we do not have a strong theory on selecting these study characteristics. Previous studies typically select a specification and infer the effects of chosen study characteristics without explicitly considering the uncertainty of the model selection procedure. Technically speaking, such a practice can understate the degree of uncertainty of inferences. To address this issue, we adopt the Bayesian Model Averaging (BMA) approach that

[12] In the next Section, an alternative set of benchmark types is considered.

explicitly accounts for both model selection and sampling uncertainties in drawing inferences on parameters of interest.

In essence, the BMA approach estimates the full posterior distribution of a parameter of interest as a weighted average of its posterior probabilities conditional on all model in the model space with weights given by the corresponding posterior model probabilities. The estimation uses information on the prior distribution of the parameter on every model on the model space, and the prior distributions of models on the model space, and the sample likelihood function. The posterior inclusion probability (PIP) of a variable is given by the sum of the posterior probabilities of models that include the variable, and is used to determine whether the variable should be included in the regression or not. Further, based on a parameter's posterior distribution, we obtain its posterior mean and posterior standard error.[13] See Appendix A.4 for a discussion of the BMA methodology.

3.1 **Basic BMA Results**

To assess which one of the 26 study characteristic types (moderator variables) is favored by the data for inclusion in (1), we assume two conservative and commonly used priors; namely, the uniform prior probability on the model space (that has 2^{26} elements) and the unit information prior g-UIP for parameters (Zeugner and Feldkircher, 2015). A Markov Chain Monte Carlo method based on the Metropolis-Hasting algorithm is employed to conduct the BMA analysis. For all BMA computations we use 1,000,000 burn-ins and 2,000,000 iterations to ensure a good degree of convergence. In passing, we note that the extreme values visualized in the box plots in the previous section are not excluded from the exercise. Instead, we let the data to determine their relevance via posterior probabilities based on the priors and the likelihood function.

Figure 7 presents information of the top 6,000 model specifications that have the highest posterior model probabilities.[14] The 26 study characteristic types (moderator variables) are listed on the vertical axis in descending orders of their posterior inclusion probabilities. Each column represents a model specification with the column width indicates its posterior model probability, which measures the degree it is favored by data. For each column, a blue cell (darker color in grayscale) implies that the corresponding study characteristic type listed on the vertical axis is

[13] For comparison purposes, the posterior inclusion probability analogizes the p-value, posterior mean the estimate, and posterior standard error the standard error under the frequentist approach.

[14] The Bayesian model sampling package in R and the Metropolis-Hasting algorithm were used to select models with high posterior model probabilities (Zeugner and Feldkircher, 2015).

included in the model specification and has a positive coefficient estimate, a red cell (lighter color in grayscale) implies the corresponding study characteristic type is included and has a negative coefficient estimate, and a blank cell means that the study characteristic type is not included in the model specification. These model specifications are presented from left to right according to their posterior model probabilities from high to low, and the cumulative posterior model probabilities are listed on the horizontal axis. The 6,000 models with highest posterior model probabilities account for about 90% of the probability on the model space.

Two study characteristic types, namely "2011-2014" and "Cross-sectional" are included in these top 6,000 model specifications – with the former characteristic type displays a consistently negative sign (red cell; lighter color in grayscale) and the latter one a consistently positive sign (blue cell; darker color in grayscale) in these specifications. Recall that the estimated effect of a given characteristic type is a marginal effect relative to that of the corresponding benchmark reference characteristic type.

The remaining study characteristic types have a declining frequency of occurrence in these top model specifications, and some of them (e.g. characteristic types labeled "FEERs" and "Cointegration") even garner coefficient estimates with different signs across specifications. It is noted that the effects of these study characteristic types are not necessarily in accordance with the size of their medians relative to their corresponding benchmark characteristic types; indicating that these study characteristic types can be correlated and, hence display effects in a multivariate framework that are different from the descriptive statistics depicted in box plots.

Table 1 presents, for each study characteristic type, some statistics derived from its full posterior probability of the parameter and the posterior model probability. Under the column label "PIP," we report the posterior inclusion probability which measures the likelihood of including a parameter in the regression. Following Kass and Raftery (1995) and Havranek *et al*. (2015), a study characteristic type is considered to have an acceptable, substantial, strong and decisive effect if it has a posterior inclusion probability falls between 0.5-0.75, 0.75-0.95, 0.95-0.99 and 0.99-1, respectively. If the posterior inclusion probability is less than 0.5, the study characteristic type is considered to be ignorable. In Table 1, a bold-type font indicates that the study characteristic type has an estimated posterior inclusion probability larger than 0.5.

The columns labeled “Post Mean” and “Post SD” report the mean and standard error computed from the full posterior distribution of a parameter, which incorporates uncertainties attributable to both model selection and sampling processes.

The column labeled “Sign” presents, based on data under examination, the confidence about the sign of the parameter of a study characteristic type (which is reflected by the color intensity of the study characteristic type row in Figure 7). Specifically, a value of one implies the parameter is positive (that is, the study characteristic type has a positive impact on RMB misalignment estimates), and a value of zero implies the parameter is negative (that is, the study characteristic type tends to reduce the size of RMB misalignment estimates). If the value is closer to one (zero), then the effect of the study characteristic type is more likely to be positive (negative).

For the frequency characteristic under the category of data characteristics, the parameter estimates of the monthly and quarterly characteristics have posterior inclusion probability values noticeably below the 0.5 threshold, and posterior means that are close to zero and small compared with their corresponding posterior standard errors. That is, based on the data, the BMA results suggest that studies using monthly and quarterly data do not yield RMB misalignment estimates that are, *ceteris paribus*, significantly different from those based on annual observations.

Results of the other two data characteristics, however, yield evidence of heterogeneity of RMB misalignment estimates across study characteristics. The “PPP-based” characteristic type has a posterior inclusion probability value of 0.954 (very close to one) and a value of 1 under the “Sign” column. Further, its posterior mean to posterior standard error ratio is 2.826. These results strongly suggest that the use of PPP-based data is likely to generate large RMB misalignment estimates; that is, strong evidence of RMB undervaluation. Note that explicit efforts were devoted to compare the cross-country purchase power parity in compiling PPP-based real exchange rate data. It is believed that China’s productivity growth underlying its dramatic economic performance is not properly reflected in the market exchange rates and prices. To the extent that relative productivity growth has implications for purchasing power and, hence, real exchange rate, the use of market exchange rates and prices is likely to understate the positive effect of productivity on real exchange rate.

On the choice of exchange rate data, the use of either data on the RMB real effective exchange rate or on other types of exchange rate including nominal (effective) exchange rates, on the other hand, tend to yield weak evidence of RMB undervaluation. These two exchange rate characteristic types have a value of zero under "Sign" and have posterior inclusion probability values larger than 0.5, which suggest that they should be included in the regression. An effective exchange rate comprises a country's exchange rates against a group of countries and is deemed to be a better measure of a country's competitiveness than a bilateral exchange rate. The BMA result suggests that the use of a bilateral real RMB-US dollar exchange rate tends to overstate the RMB's general level of undervaluation.

The BMA results indicate that the theoretical frameworks considered in this meta-analysis do not contribute to the cross-study heterogeneity of misalignment estimates. The four types "FEERs," "Penn effect," "PPP," and "Other frameworks" have a posterior inclusion probability value less than 0.5, and a very small posterior mean relative to the posterior standard error. That is, these theoretical frameworks is not likely to generate misalignment estimates different from those derived from the "BEERs" specification. Despite the discussions of advantages and disadvantages of different theoretical frameworks (Clark and MacDonald, 1999; Lopez-Villavicencio *et al.*, 2012), our results do not show RMB misalignment estimates are systematically affected by the choice of theoretical frameworks underlying the empirical exercise. Further, the use of either cointegration or non-cointegration techniques is not likely to be the source of study-to-study variations of misalignment estimates as the "Cointegration" type in Table 1 has a small posterior inclusion probability of 0.108.

Under the category of theoretical and estimation specifications, the "Cross-sectional" and "Structural" are the two study characteristic types that have posterior inclusion probability values larger than 0.9 and exhibit positive effects on the misalignment estimates. Under a time series model specification, the (estimated) equilibrium exchange rate is typically given by the average over time (conditional on regressors) and, hence, overvaluation and undervaluation estimates are almost invariably to be reported. When cross-sectional data are used, it is implicitly assumed that, conditional on regressors, the sample average across countries is the (estimated) equilibrium exchange rate. Thus, the use of cross-sectional data allows for the possibility that the RMB is, say, undervalued, for an extended period and is not at the (estimated) equilibrium value. The significance of the "Cross-sectional" characteristic type is suggestive of the estimated equilibrium exchange rate based on cross-country averages is larger than the RMB rate.

Our significant “Structural” characteristic type result is comparable to Wang and Yao (2008), who find that structural setup usually give larger misalignment estimates relative to reduced-form approach.

The “Mixed” characteristic type under the category of publication attributes is the only characteristic type that has a posterior inclusion probability estimate larger than 0.5. The “Mixed” characteristic type refers to studies that have authors from more than one type of these institutions namely academics, government, and industry. When authors are from different types of institutions, they tend to present a relatively weak evidence of RMB undervaluation.

The BMA results show that the RMB misalignment estimates generated for the periods of 2009 to 2010 and of 2011 to 2014 tend to be different from those for other periods. Posterior inclusion probabilities indicate that, compared with the period 2009 to 2010, estimates generated for assessing RMB misalignment in 2011 to 2014 should have a higher chance to affect misalignment estimates. While both cases have a value of zero under the “Sign” column; indicating a negative impact is observed, the 2011 to 2014 misalignment estimates are more likely to display a small undervaluation value. The finding is in accordance with the anecdotal evidence of a strengthening RMB and narrowing Chinese current account surplus observed after the Global Finance Crisis.

The meta-analysis based on the BMA method offers evidence of the reported RMB misalignment estimates are associated with some study characteristics including the property of the data used, the choice of theoretical and estimation methods, author’s affiliations, and the periods in which the estimates are generated for. The empirical effects of study characteristics have been accounted for not only sampling variations but also uncertainties related to the choice of model specifications.

3.2 Misalignment under Hypothetical Combinations of Study Characteristics

From Table 1, we identify eight study characteristic types that have a value of posterior inclusion probability larger than 0.5. These eight variables constitute the so-called "Median Probability Model."[15]

[15] Note that in the current study, the median probability model and the highest probability model, which is the model specification that has the highest posterior model probability (Barbieri and Berger, 2004), are the same because both models include the same set of eight study characteristic types (Figure 7).

What can we say about the RMB misalignment estimate when it is generated from a study with the eight study characteristic types under the median probability model? If we assume these variables take up their respective sample average values, then the resulting RMB misalignment estimate is the average of what are reported in our sample conditional on these study characteristic types. Barbieri and Berger (2004), for instance, indicates that the median probability model yields good predictions.

Adopting the median probability model specification, we obtain the density plot of the RMB misalignment estimate as depicted in Figure 8. The 2.5%, 50% and 97.5% quantiles of the estimated density are, respectively, -0.363, 0.116 and 0.596. That is, while the median of the estimate is positive, the 95% confidence interval however indicates the estimate is not significantly different from zero. We also consider the case in which the eight variables assume the value of one; that is, a study that possesses these eight study characteristic types.[16] The resulting 2.5%, 50% and 97.5% quantiles of the RMB misalignment estimate are, respectively, -0.741, -0.182 and 0.392. The change of the assumed values of study characteristic types decreases the median of the RMB misalignment estimate from positive to negative; that is, the evidence leans towards RMB overvaluation. Nevertheless, the estimate is not significantly different from zero.

Next, we split the eight variables of the median probability model into two groups; one comprises three study characteristic types with a positive sign, and one five study characteristic types with a negative sign. When the three study characteristic types with a positive sign assume the value of one (and the others a value of zero), the 2.5%, 50% and 97.5% quantiles of the RMB misalignment estimate are -0.022, 0.477, and 0.969. For the case of five study characteristic types with a negative sign, the 2.5%, 50% and 97.5% quantiles are -1.114, -0.571 and -0.015. That is, if a hypothetical study is conducted with the three identified "positive" study characteristic types, there is no more than 95% "confidence" of obtaining an undervalued RMB inference. On the other hand, if the hypothetical study is equipped with the five identified "negative" study characteristic types, then, there is more than 95% chance that the RMB is found to be overvalued.

The average model approach that considers all the study characteristic types (moderator variables) is another way to generate the prediction of RMB misalignment (Eklund and Karlsson,

---

[16] Strictly speaking, this setup is not feasible because it includes both "REER" and "NER/NEER" and both "2009-2010" and "2011-2014". The same insignificant result is obtained when alternative combinations of study characteristic types with a feasible subsample configuration are considered.

2007; Feldkircher, 2012). Specifically, the average model approach sets all the study characteristic types to their respective average values to generate the corresponding information of RMB misalignment. The symmetric 95% confidence interval around the posterior median of the RMB misalignment estimate is (-0.342, 0.135, 0.611); indicating the estimate is insignificantly different from zero.

We consider a few other hypothetical combinations of study characteristic types including a) one that includes the least commonly used study characteristic types, and b) those consider different subsample periods. In all these cases, we obtained similar insignificant results – the 95% confidence interval of the RMB misalignment estimate includes both negative and positive values.[17]

One interpretation of these results from hypothetical combinations of study characteristics is that the information embedded in these studies of RMB misalignment is quite diverse. With the exception of the case of a hypothetical study that is equipped with the five identified negative characteristic types, it is hard to draw a definitive inference to reject the hypothesis of the RMB is not misaligned once we have controlled for model selection and estimation uncertainties. It is of interest to note that, for example, Cheung *et al.* (2007), Dunaway *et al.* (2009) and Schnatz (2011) argue that the data are not sufficiently informative to give a clear-cut inference about RMB misalignment based on different non-Bayesian settings.

4. Additional Analyses

To shed some additional insights on the effects of study characteristics on the reported RMB misalignment estimates, we present results from the least squares approach and from using an alternative benchmark specification.

4.1 Regression Analysis

In this subsection, we present results based on least squares estimation. The least squares result of estimating the median probability model specification identified in the previous section is presented in Table 2 under column 1. All the eight study characteristic types of the median probability model garner a statistically significant coefficient estimate with a sign that is consistent with the one revealed in Table 1. The magnitudes of these least squares estimates, however, can be

[17] Results of these cases are available from the authors; also see footnote 18.

quite different from the corresponding posterior means in Table 1. Apparently, the difference between the least squares estimate and the corresponding posterior mean is inversely related to the posterior inclusion probability of the study characteristic type. The difference is relatively small for the study characteristic types such as “PPP-based,” “Cross-sectional,” “Structural,” and “2011-2014” that have a posterior inclusion probability larger than 90%. Recall that under the BMA approach, the posterior inclusion probability is used to infer whether a variable is expected to have a substantial impact on the RMB misalignment estimate.

The least squares coefficient estimates of the 26 study characteristic types are presented under column 2 in Table 2. There are 19 statistically significant study characteristic types. Among the eight study characteristic types included in the median probability model specification, the “PPP-based” type is the only one that is insignificant. The least squares approach which does not explicitly consider the uncertainty of model selection indicates the magnitude of RMB misalignment is affected by a large number of characteristic types. Despite the noticeable increase in the number of significant explanatory variables, the improvement in model performance does not appear substantial. As indicated by the adjusted R^2 estimate, the 26-variable specification explains 25.7% of the variability of RMB misalignment estimates; the explanatory power is slightly higher than the 24.3% offered by the 8-variable median probability model specification.

A possible reason of the least squares approach overstating the significance of study characteristics is that it does not account for the possibility that RMB misalignment estimates generated by a study can be correlated even though they are independent across studies. If RMB misalignment estimates cluster on study, then the usual heterokedasticity-robust standard errors reported in columns (1) and (2) can greatly overstate estimation precision, and yield spurious significant results. To entertain this possibility, columns (3) and (4) present standard errors that are robust to clustering on study.

Arguably, the statistical inference of the 26-variable specification, relative to the 8-variable median probability model, is greatly affected by the use of cluster-robust standard errors. Specifically, when cluster-robust standard errors are used to infer statistical significance, the number of significant study characteristic types of the former model drops to 7 from 19 while the latter one to 7 from 8. That is, the inference based on the median probability model specification is quite robust to the clustering on study, and the standard least squares results can exaggerate the impact of study characteristics on the reported RMB misalignment estimates.

Based on cluster-robust standard errors, both model specifications have 7 significant explanatory variables; however, only three significant study characteristic types (“BEERs,” NER/NEER,” “2011-2014”) are common to these two specifications. That is, study characteristic types identified by the BMA are different from those by the least squares method with or without controlling for clustering effects.

4.2. Alternative BMA Results

In the previous Section, the effects of study characteristics were evaluated relative to the selected benchmarks. What will happen if different benchmarks are selected - for example, the quarterly and monthly data instead of annual data are the benchmark characteristic type of the data frequency study characteristic? In this sub-section, we re-assess the results when we have inter-changed the roles of benchmark and non-benchmark characteristic types used in the previous section. That is, the 26 study characteristic types considered in the previous section assume the role of benchmark types of their corresponding study characteristics. At the same time, the 13 benchmark types of the previous section (those in bold font in Table A.3) become the regressors of equation (1).

Under this arrangement, there are 2^{13} (= 8,192) possible empirical specifications. Figure 9 and Table 3 presents the corresponding BMA results. Using the format of Figure 7, Figure 9 presents information of the top 100 model specifications that have the highest posterior model probabilities, and they together account for about 90% of the probability on the model space.

The specification that has three study characteristic types of “Time series,” “Chinese,” and “BEERs” garners the highest posterior model probability of 0.14. For this specification, the use of time-series data or a BEERs framework is likely to yield a small RMB misalignment estimate,[18] and a study that has an author with a Chinese name is likely to generate a large RMB misalignment estimate. In addition to these three types, the characteristic type “Reduced-form” also garners a posterior inclusion probability larger than 0.5 (Table 3). That is, for this group of 13 study characteristic types, the median probability model has four explanatory variables; namely, “Time series,” “Chinese,” “BEERs,” and “Reduced-form”.

When the four explanatory variables of the median probability model are set to their sample average values (and the remaining nine variables set to zero), the 2.5%, 50% and 97.5% quantiles

[18] Similar negative “BEERs” and “Time series” effects are reported in Korhonen and Ritola (2011).

of the distribution of the resulting RMB misalignment estimate are given by -0.387, 0.130, and 0.647. When the four variables assume the value of one, then 2.5%, 50% and 97.5% quantiles become -0.473, 0.045, and 0.563. In both cases, based on the 95% confidence interval, the misalignment estimate is insignificantly different from zero. A similar insignificant result is obtained under the average model framework in which all the 13 study characteristic types assume their sample average values, or the value of one.[19]

Table 4 presents the least squares regression results of the median probability model and average model specifications. Column (1) presents the estimates of the median probability model specification and their corresponding robust standard errors. The study characteristic "Reduced-form," which has the smallest posterior inclusion probability among the four explanatory variables is not significant. The other three study characteristic types are statistically significant with signs in accordance with those indicated in Table 3. When all the 13 variables are included in the least squares regression, 12 of them are significant (column (2)).

The standard errors that are robust to clustering on study are presented under columns (3) and (4). Controlling for clustering effects noticeably affects the pattern of significant study characteristics. Specifically, after controlling for clustering effects, the median probability model specification yields one significant characteristic type ("Time series") while the 13-variable specification yields two significant characteristic types ("Time series" and "Chinese"). Similar to findings from Table 2, the generic least squares approach overstates the number of significant study characteristics, and the spurious significant result is likely attributable to clustering effects.

In comparing results in the current and previous sections, it is clear that, in addition to the choice of the BMA and least squares approach, the choice of study characteristic benchmarks affects the empirical significance of study characteristics. Apparently, the "Time series" or the non-time series "Cross-sectional" is the only study characteristic that is significant in all these different specifications.

The finding is likely attributable to the artifact that, in general, a time series specification typically identifies the equilibrium exchange rate with the average (conditional on regressors) over the sample period; that is, there are both overvaluation and undervaluation estimates in the sample and, on average, the exchange rate is at equilibrium. A cross-sectional model, on the other hand,

[19] The (2.5%, 50%, 97.5%) quantiles of the misalignment estimate distribution are, respectively, (-0.372, 0.135, 0.641) and (-0.467, 0.047, 0.561).

typically identifies the equilibrium exchange rate with the average (conditional on regressors) across countries in the sample. If, over time, the RMB is consistently undervalued relative to the equilibrium value defined by the cross-country average, then the RMB can be reported as undervalued for an extended period and is not at the (estimated) equilibrium value in the cross-sectional setting.

In passing, it is noted that the adjusted R^2 estimates indicate that specifications in Table 2 offer better a better explanatory power than the corresponding ones in Table 4.

5. Concluding Remarks

We adopt the BMA approach to conduct a meta-analysis of the effect of study characteristics on empirical RMB misalignment estimates. One advantage of the BMA framework is that it explicitly accounts for the uncertainties of model selection and sampling in assessing the effects of study characteristics. Our exercise includes 937 RMB misalignment estimates obtained from 69 studies, and 13 study characteristics.

The basic BMA results show that the study-to-study heterogeneity of RMB misalignment estimates is associated with some of the selected study characteristics. However, the result is sensitive to the estimation technique - the least squares method yields a different set of significant study characteristics. Also, a different choice of study characteristic benchmarks generates different significant study characteristics. The only study characteristic that is found to be significant in all the reported results is the use of a time series specification (or a cross-sectional specification).

In assessing the uncertainty of the RMB misalignment, we derive the distribution of the RMB misalignment estimate from hypothetical combinations of study characteristics; including the median probability model and average model obtained under the BMA approach. It is found that, other than one exception of all the hypothetical cases we have considered, the RMB misalignment estimate is insignificantly different from zero; that is, we cannot reject the notion of the RMB is not misalignment.

We employ the meta-analysis to conduct a formal quantitative analysis of studies on RMB misalignment. If these studies are based on some common true conceptual model of the RMB, then the study-to-study variability of misalignment estimates can be attributed to random (measurement) errors within or across individual studies. The pooling under the meta-analysis setup shall aggregate information and lead to a precise estimate. Apparently, we do not obtain a definite evidence on the heterogeneity of RMB misalignment estimates. Our results show that the significance of study characteristics varies quite substantially across the Bayesian and classical least squares approaches, and different choices of benchmark characteristics. Further, the evidence of a misaligned RMB is quite weak.

One interpretation of our findings is that we should be very cautious in assessing and interpreting the reported RMB misalignment estimates. Cheung *et al*. (2007), Dunaway *et al*. (2009) and Schnatz (2011), for example, have noted that the evaluation of misalignment is hindered by the absence of a consensual equilibrium exchange rate model, substantial data revisions, and the sensitivity to small changes in assumptions underlying empirical specifications. Both the theoretical and empirical data are not sufficiently informative for deciphering the equilibrium value and, hence, the degree of misalignment. These factors together can prevent our exercise to give a precise inference about the study characteristic effects and the RMB misalignment.

References

Agya, Atabani Adi, and Jun, Du (2015). Estimate of equilibrium real exchange rate and misalignment of Chinese Yuan vis-a-vis US dollar. Academic Journal of Economic Studies, 1(3), 148-165.

Almas, Ingvild, Grewal, Mandeep, Hvide, Marielle, and Ugurlu, Serhat (2017). The PPP approach revisited: A study of RMB valuation against the USD. Journal of International Money and Finance, 77, 18-38.

Barbieri, Maria Maddalena, and Berger, James O. (2004). Optimal predictive model selection. The Annals of Statistics, 32(3), 870-897.

Bineau, Yannick (2010). Renminbi's misalignment: A meta-analysis. Economic Systems, 34(3), 259-269.

Chang, Gene Hsin, and Qin, Shao (2004). How much is the Chinese currency undervalued? A quantitative estimation. China Economic Review, 15(3), 366-371.

Cheung, Yin-Wong, Chinn, Menzie D., and Fujii, Eiji (2007). The overvaluation of renminbi undervaluation. Journal of International Money and Finance, 26(5), 762-785.

Cheung, Yin-Wong, Chinn, Menzie D., and Nong Xin (2017). Estimating currency misalignment using the Penn effect: It is not as simple as it looks. International Finance, 20(3), 222-242.

Cheung, Yin-Wong, Chinn, Menzie D., and Pascual, Antonio Garcia (2005). Empirical exchange rate models of the nineties: Are any fit to survive? Journal of International Money and Finance, 24(7), 1150-1175.

Cheung, Yin-Wong, Chinn, Menzie D., and Pascual, Antonio Garcia, and Zhang, Yi (2018). Exchange rate prediction redux: New models, new data, new currencies. Journal of International Money and Finance. (Forthcoming, https://doi.org/10.1016/j.jimonfin.2018.03.010)

Chou, W. L., and Shih, Y. (1998). The equilibrium exchange rate of the Chinese Renminbi. Journal of Comparative Economics, 26(1), 165-174.

Clark, Peter B., and MacDonald, Ronald (1999). "Exchange rates and economic fundamentals: A methodological comparison of BEERs and FEERs Equilibrium exchange rates," in Ronald MacDonald and Jerome L. Stein (Eds.), Equilibrium Exchange Rates (Recent Economic Thought Series, Volume 69, Chapter 10, pp. 285-322). Dordrecht: Springer.

Dunaway, Steven, Leigh, Lamin, and Li, Xiangming (2009). How robust are estimates of equilibrium real exchange rates: The case of China. Pacific Economic Review, 14(3), 361-375.

Égert, Bala´zs, and Halpern, La´szlo´ (2006). Equilibrium exchange rates in Central and Eastern Europe: A meta-regression analysis. Journal of Banking & Finance, 30(5), 1359-1374.

Eklund, Jana, and Karlsson, Sune (2007). An embarrassment of riches: Forecasting using large panels, Working Papers 2007: 1.

Feldkircher, Martin (2012). Forecast combination and Bayesian model averaging: A prior sensitivity analysis. Journal of Forecasting, 31(4), 361-376.

Funke, Michael, and Rahn, Jörg (2005). Just how undervalued is the Chinese renminbi? The World Economy, 28(4), 465-489.

Giannellis, Nikolaos, and Koukouritakis, Minoas (2018). Currency Misalignments in the BRIICS Countries: Fixed Vs. Floating Exchange Rates. Open Economies Review, 1-29.

Havranek, Tomas, Horvath, Roman, Irsova, Zuzana, and Rusnak, Marek (2015). Cross-country heterogeneity in intertemporal substitution. Journal of International Economics, 96(1), 100-118.

Meese, Richard A., and Rogoff, Kenneth (1983). Do they fit out of sample? Journal of International Economics, 14, 3-24.

Kass, Robert E., and Raftery, Adrian E. (1995). Bayes factors. Journal of the American Statistical Association, 90(430), 773-795.

Korhonen, Iikka, and Maria Ritola, (2011), "Renminbi Misaligned — Results from Meta-Regressions," in Yin-Wong Cheung and Guonan Ma (Eds.), Asia and China in the Global Economy (Chapter 4, pp. 97–122). Singapore: World Scientific Publishing Company.

Li, Chang (2015). Equilibrium Exchange Rate and Misalignment of the Chinese Renminbi: A Research Based on Montiel's Model and the BEER Approach (Master's thesis).

Lopez-Villavicencio, Antonia, Mazier, Jacques, and Saadaoui, Jamel (2012). Temporal dimension and equilibrium exchange rate: A FEER/BEER comparison. Emerging Markets Review, 13(1), 58-77.

Rossi, Barbara (2013). Exchange rate predictability. Journal of Economic Literature, 51(4), 1063-1119.

Schnatz, Bernd (2011). Global imbalances and the pretence of knowing fundamental equilibrium exchange rates. Pacific Economic Review, 16(5), 604-615.

Tukey, John W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley.

Zeugner, Stefan, and Feldkircher, Martin (2015). Bayesian model averaging employing fixed and flexible priors: The BMS package for R. Journal of Statistical Software, 68(4), 1-37.

Zhang, Zhichao (2001). Real exchange rate misalignment in China: An empirical investigation. Journal of Comparative Economics, 29(1), 80-94.

Wang Bing 王彬. (2015). Renminbi exchange rate equilibrium, imbalance and trade surplus adjustment (人民币汇率均衡、失衡与贸易顺差调整). Jingjixue(jikan)(04), 1277-1302. Available from CNKI

Wang Zetian 王泽填, and Yao Yang 姚洋. (2008). Estimating RMB equilibrium exchange rate (人民币均衡汇率估计). Jinrongyanjiu (12), 22-36. Available from CNKI

Zhang Xiaopu 张晓朴. (2000). Equilibrium and misalignment: the assessment of RMB exchange rate during 1978-1999 (均衡与失调：1978-1999 人民币汇率合理性评估). Jingjiyanjiu (08), 13-24. Available from CNKI

Table 1. The BMA results

	PIP	Post Mean	Post SD	Sign
a) Data Characteristics				
Quarterly	0.258	-0.017	0.033	0.000
Monthly	0.044	0.004	0.027	0.973
PPP-based	**0.954**	**0.130**	**0.046**	**1.000**
REER	**0.597**	**-0.048**	**0.044**	**0.000**
Other RERs	0.031	0.000	0.010	0.630
NER/NEER	**0.850**	**-0.148**	**0.080**	**0.000**
b) Theoretical and Estimation Specifications				
FEERs	0.082	0.006	0.029	0.798
Penn effect	0.296	0.033	0.059	1.000
PPP	0.064	-0.003	0.021	0.106
Other frameworks	0.062	0.003	0.019	0.729
Panel	0.302	0.023	0.040	0.999
Cross-sectional	**0.999**	**0.154**	**0.046**	**1.000**
Cointegration	0.108	0.006	0.024	0.932
Structural	**0.943**	**0.105**	**0.039**	**1.000**
c) The Publication Attributes				
Book chapter	0.036	0.002	0.019	1.000
Other types	0.043	0.001	0.008	0.978
Chinese study	0.109	-0.011	0.038	0.000
Mainland	0.064	-0.002	0.009	0.011
Non-Chinese	0.046	0.001	0.007	0.761
Government	0.418	0.028	0.038	1.000
Industry	0.028	-0.001	0.042	0.178
Mixed	**0.688**	**-0.074**	**0.058**	**0.000**
d) Subsample Periods				
1994-1997	0.030	0.000	0.004	0.735
2005-2008	0.038	0.001	0.005	0.986
2009-2010	**0.672**	**-0.067**	**0.055**	**0.000**
2011-2014	**1.000**	**-0.319**	**0.032**	**0.000**
(Intercept)	1.000	0.087	NA	NA

Notes: A bold font indicates that the study characteristic type has an estimated posterior inclusion probability larger than 0.5. “PIP” refers to posterior inclusion probability which measures the likelihood of including a parameter in the regression. “Post Mean” and “Post SD” report the mean and standard error computed from the full posterior distribution of a parameter. “Sign” presents the confidence about the sign of the parameter; the closer to “1” (“0”) the more likely the effect of the corresponding study characteristic type is positive (negative).

Table 2. OLS results-robustness check - I

	1	2	3	4
a) Data Characteristics				
Quarterly		-0.068**		-0.068
		(0.028)		(0.063)
Monthly		0.116*		0.116
		(0.065)		(0.082)
PPP-based	0.149***	0.034	0.149***	0.034
	(0.023)	(0.036)	(0.054)	(0.070)
REER	-0.086***	-0.076***	-0.086*	-0.076*
	(0.021)	(0.021)	(0.045)	(0.044)
Other RERs		-0.013		-0.013
		(0.064)		(0.074)
NER/NEER	-0.219***	-0.188***	-0.219***	-0.188**
	(0.028)	(0.048)	(0.037)	(0.073)
b) Theoretical and Estimation Specifications				
FEERs		-0.028		-0.028
		(0.092)		(0.178)
Penn effect		0.245***		0.245**
		(0.053)		(0.105)
PPP		0.220***		0.220
		(0.080)		(0.148)
Other frameworks		0.182**		0.182
		(0.081)		(0.143)
Panel		0.107**		0.107
		(0.042)		(0.088)
Cross-sectional	0.141***	0.198***	0.141*	0.198
	(0.045)	(0.069)	(0.075)	(0.137)
Cointegration		0.201***		0.201*
		(0.047)		(0.102)
Structural	0.117***	0.188**	0.117***	0.188
	(0.015)	(0.085)	(0.034)	(0.147)
c) The Publication Attributes				
Book chapter		0.003		0.003
		(0.070)		(0.107)
Other types		0.042*		0.042
		(0.024)		(0.050)
Chinese study		-0.139***		-0.139
		(0.047)		(0.089)

Mainland		-0.051		-0.051
		(0.036)		(0.080)
Non-Chinese		-0.086**		-0.086
		(0.038)		(0.078)
Government		0.123***		0.123*
		(0.029)		(0.064)
Industry		0.111**		0.111
		(0.051)		(0.089)
Mixed	-0.132***	-0.066*	-0.132***	-0.066
	(0.022)	(0.036)	(0.047)	(0.072)
d) Subsample Periods				
1994-1997		0.015		0.015
		(0.028)		(0.055)
2005-2008		-0.011		-0.011
		(0.020)		(0.037)
2009-2010	-0.096***	-0.113***	-0.096	-0.113*
	(0.026)	(0.027)	(0.062)	(0.057)
2011-2014	-0.324***	-0.300***	-0.324***	-0.300***
	(0.039)	(0.040)	(0.066)	(0.065)
Constant	0.127***	-0.036	0.127**	-0.036
	(0.022)	(0.053)	(0.050)	(0.105)
Adjusted R2	0.243	0.257	0.243	0.257

Note: The lease squares regression results are presented. Columns 1 and 3 present the results of the median probability model identified by the BMA analysis, and columns 2 and 4 the results of the 26-study-characteristic-type model. The White heteroskedasticity-robust standard errors are presented in parentheses under the coefficient estimates under columns 1 and 2. The standard errors that are robust to clustering by studies are presented in parentheses under the coefficient estimates under columns 3 and 4. See the text for further detail.

Table 3. The BMA results based on 13 study characteristic types

	PIP	Post Mean	Post SD	Sign
a) Data Characteristics				
Annual	0.079	0.003	0.014	0.999
Non-PPP-based	0.470	-0.041	0.049	0.000
Dollar-based RER	0.480	0.027	0.033	1.000
b) Theoretical and Estimation Specifications				
BEERs	**0.525**	**-0.046**	**0.049**	**0.001**
Time series	**1.000**	**-0.185**	**0.030**	**0.000**
Non-cointegration	0.084	-0.004	0.020	0.126
Reduced-form	**0.506**	**-0.048**	**0.052**	**0.001**
c) The Publication Attributes				
Journal	0.082	-0.003	0.012	0.002
English study	0.048	0.003	0.018	0.973
Non-Mainland	0.322	0.024	0.040	0.989
Chinese	**0.936**	**0.091**	**0.047**	**1.000**
Academics	0.173	-0.008	0.020	0.006
d) Subsample Periods				
1998-2004	0.119	0.004	0.012	1.000
(Intercept)	1.000	0.231	NA	NA

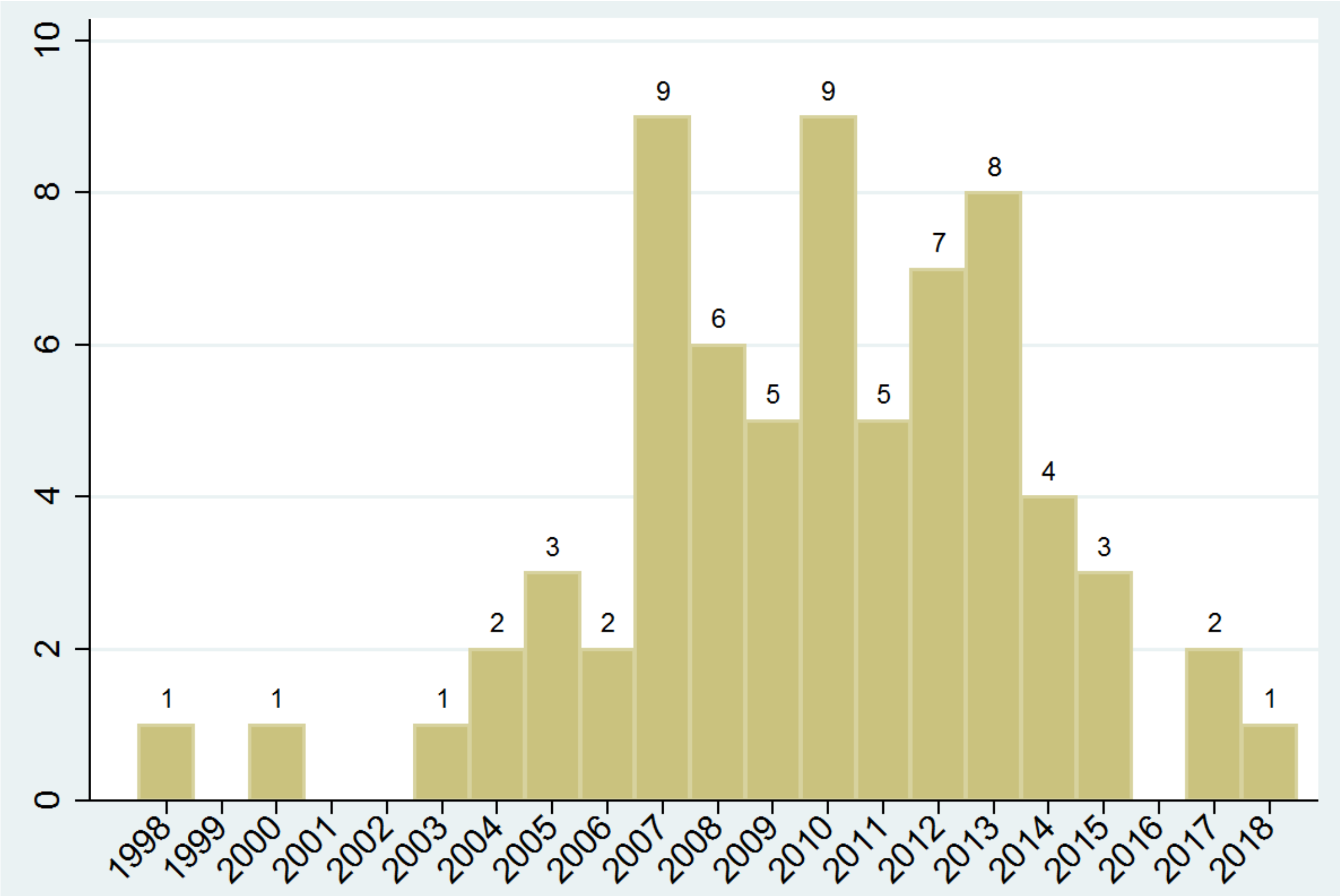
Note: See the Note to Table 1.

Table 4. OLS results-robustness check –II

	1	2	3	4
a) Data Characteristics				
Annual		0.055**		0.055
		(0.024)		(0.061)
Non-PPP-based		-0.090***		-0.090
		(0.027)		(0.063)
Dollar-based RER		0.075***		0.075
		(0.019)		(0.045)
b) Theoretical and Estimation Specifications				
BEERs	-0.074***	-0.096***	-0.074	-0.096
	(0.025)	(0.036)	(0.064)	(0.097)
Time series	-0.199***	-0.170***	-0.199***	-0.170**
	(0.025)	(0.031)	(0.061)	(0.071)
Non-cointegration		-0.091**		-0.091
		(0.038)		(0.101)
Reduced-form	-0.024	-0.071***	-0.024	-0.071
	(0.022)	(0.026)	(0.064)	(0.076)
c) The Publication Attributes				
Journal		-0.046***		-0.046
		(0.018)		(0.044)
English study		0.017		0.017
		(0.037)		(0.078)
Non-Mainland		0.113***		0.113
		(0.029)		(0.073)
Chinese	0.082***	0.176***	0.082	0.176*
	(0.029)	(0.036)	(0.079)	(0.094)
Academics		-0.072***		-0.072
		(0.020)		(0.059)
d) Subsample Periods				
1998-2004		0.038**		0.038
		(0.018)		(0.046)
Constant	0.242***	0.216***	0.242***	0.216
	(0.014)	(0.066)	(0.036)	(0.146)
Adjusted R2	0.139	0.165	0.139	0.165

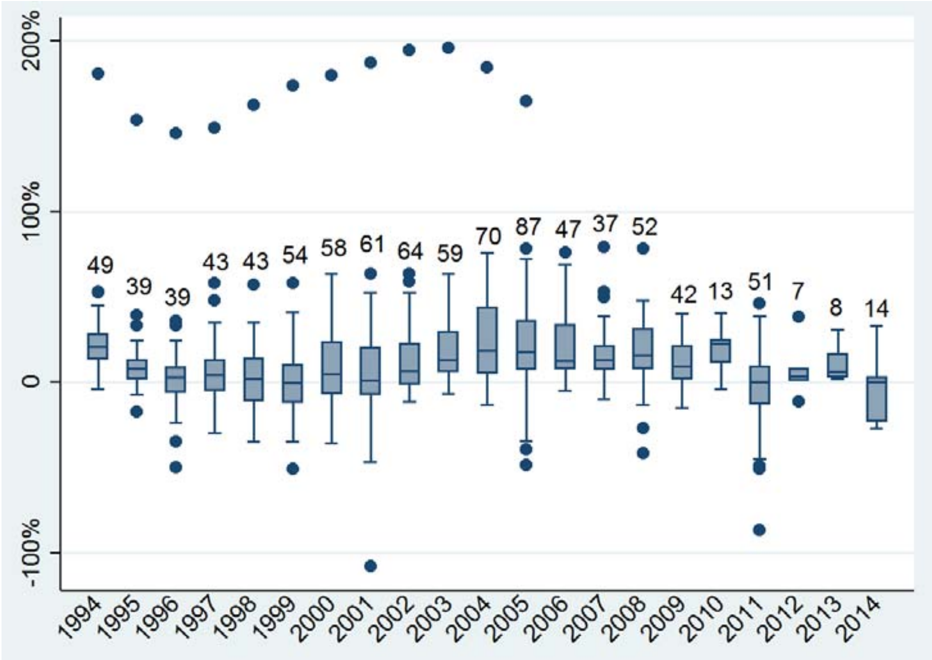
Note: See the Note to Table 2.

Figure 1. The frequency of selected studies: Year of publication



Note: The year of publication is listed on the x-axis, and the frequency of selected studies published in a given year is indicated by the bar chart.

Figure 2. The box plots of RMB misalignment estimates: Individual years



Note: The upper adjacent value = 75th percentile + (75th percentile – 25th percentile) * 1.5; the lower adjacent value = 25th percentile – (75th percentile – 25th percentile) * 1.5. The dots beyond lower/upper adjacent values are suspected outliers/extreme values. The legend of the box plot is (Tukey, 1977):

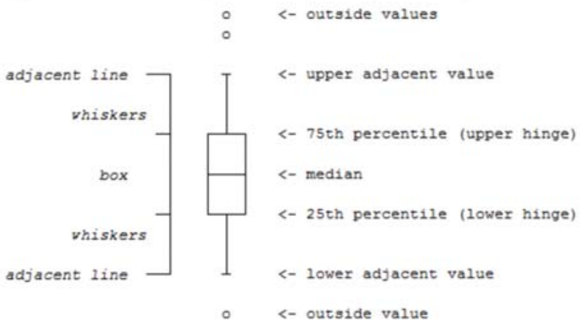
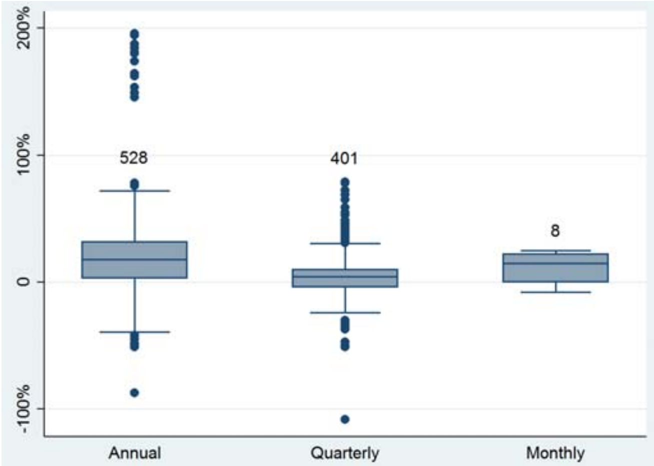
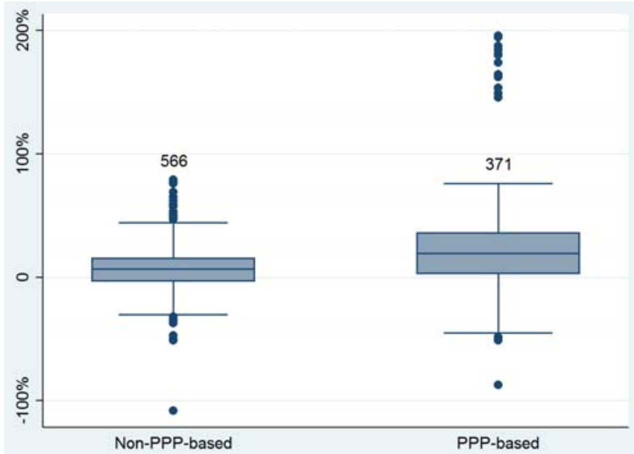


Figure 3. The box plots of RMB misalignment estimates: Data characteristics

3.a. Data frequency



3.b. Data type



3.c. Definition of the RMB exchange rate

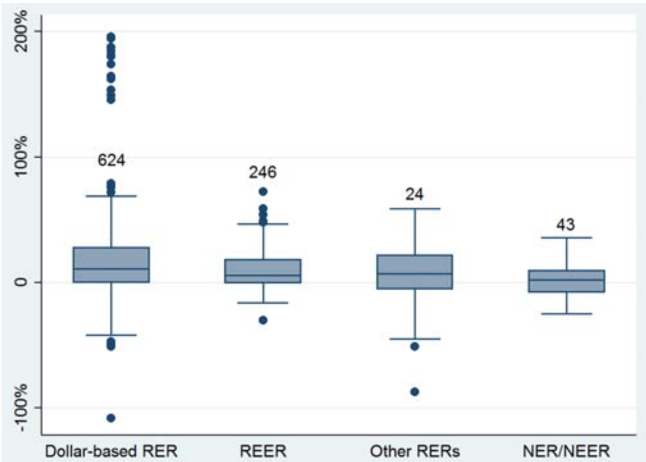
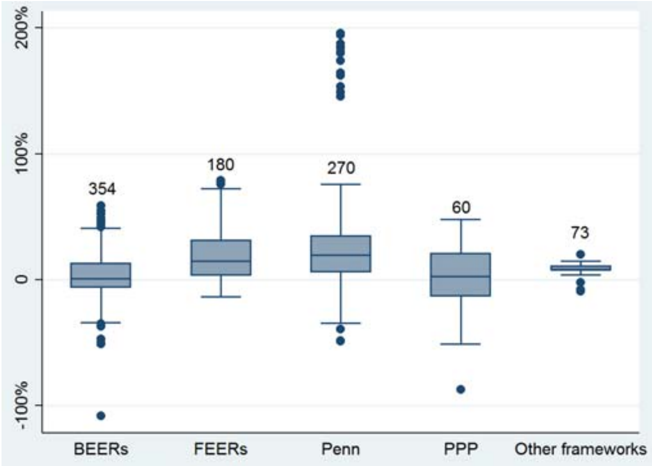
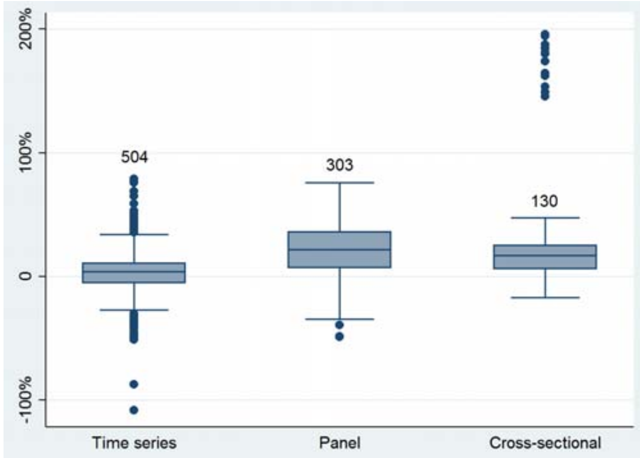


Figure 4. The box plots of RMB misalignment estimates: Theoretical and estimation Specifications

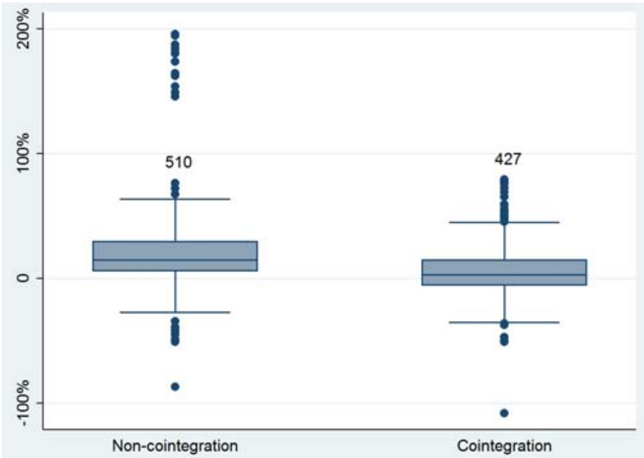
4.a. Theoretical Specifications



4.b. Estimation Method (1)



4.c. Estimation Method (2)



4.d. Estimation Method (3)

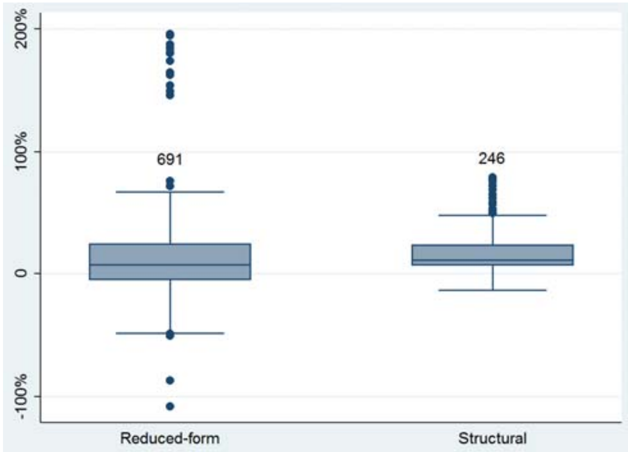
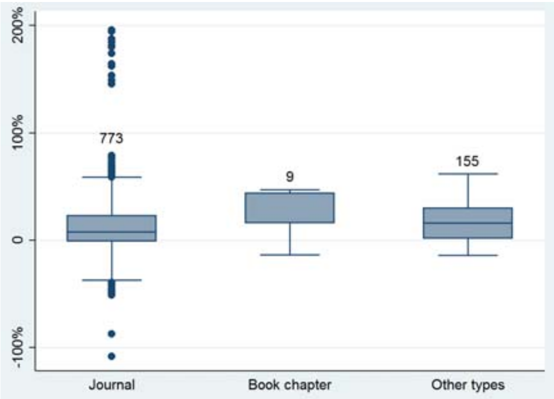
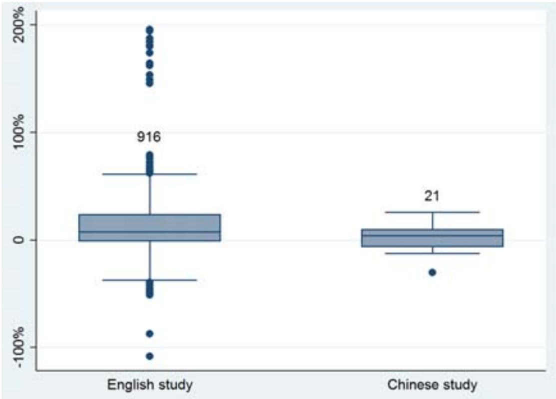


Figure 5. The box plots of RMB misalignment estimates: Publication attributes

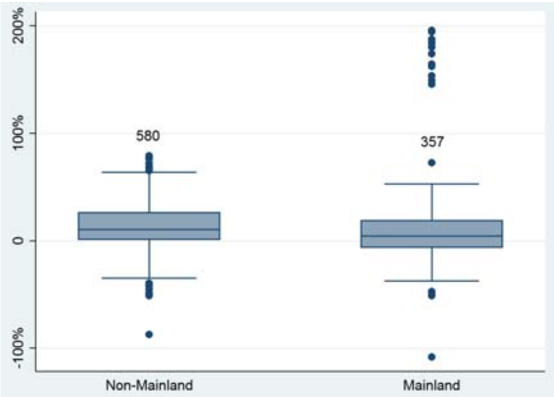
5.a Publication type



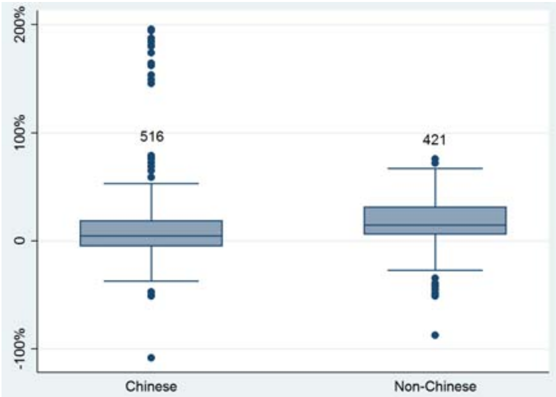
5.b Language of Publication



5.c Mainland China institutional affiliation



5.d Chinese or non-Chinese authors



5.e Types of Author's affiliations

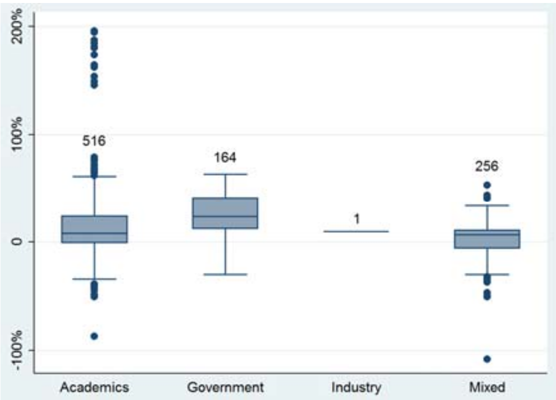


Figure 6. The box plots of RMB misalignment estimates: Subsample periods

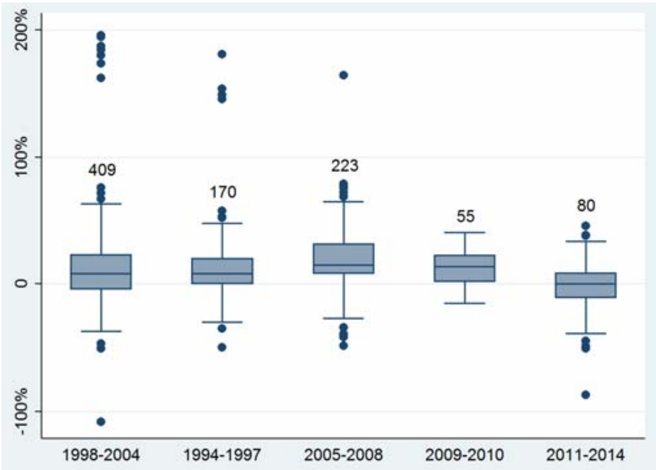
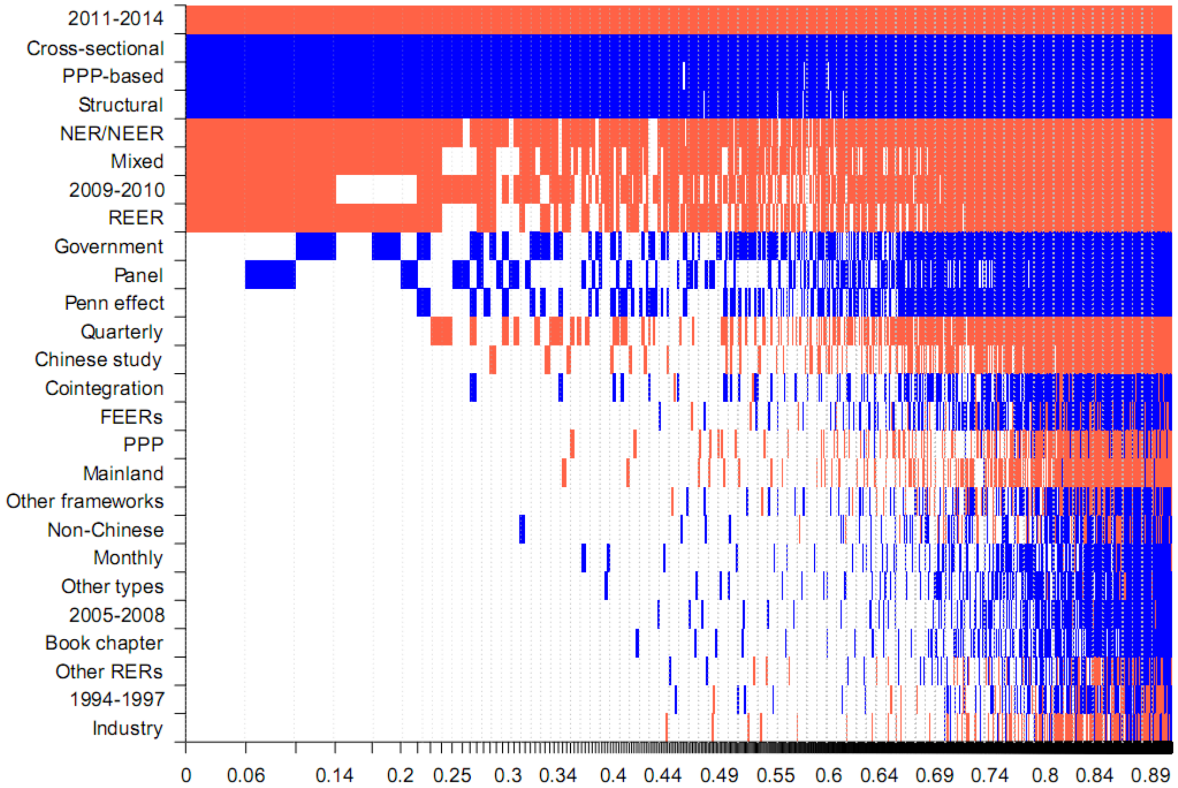


Figure 7. The top 6,000 models that have the highest posterior model probabilities (based on 26 study characteristic types)



Note: The 26 study characteristic types (moderator variables) are listed on the vertical axis in descending orders of their posterior inclusion probabilities. Each column represents a model specification with the column width indicates its posterior model probability, which measures the degree it is favored by data. For each column, a blue cell (darker color in grayscale) implies that the corresponding study characteristic type listed on the vertical axis is included in the model specification and displays a positive estimated effect, a red cell (lighter color in grayscale) implies the corresponding study characteristic type is included and displays a negative estimated effect, and a blank cell means that the study characteristic type is not included in the model specification. These model specifications are presented from left to right according to their posterior model probabilities from high to low, and the cumulative posterior model probabilities are listed on the horizontal axis.

Figure 8. The density Plot of the RMB misalignment estimate under the median probability model specification

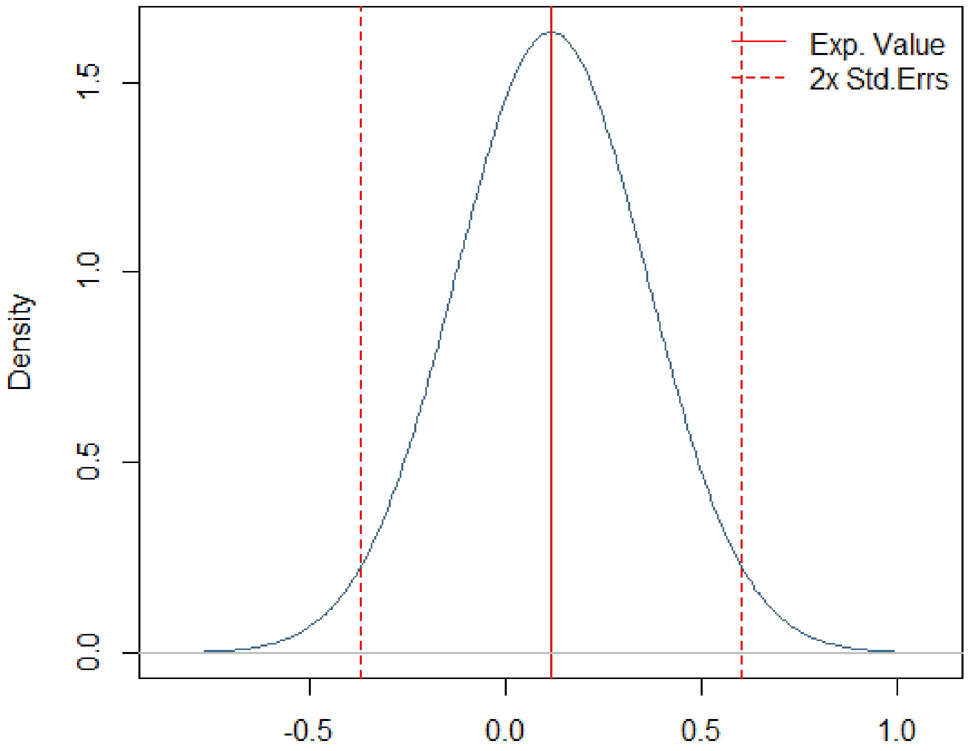
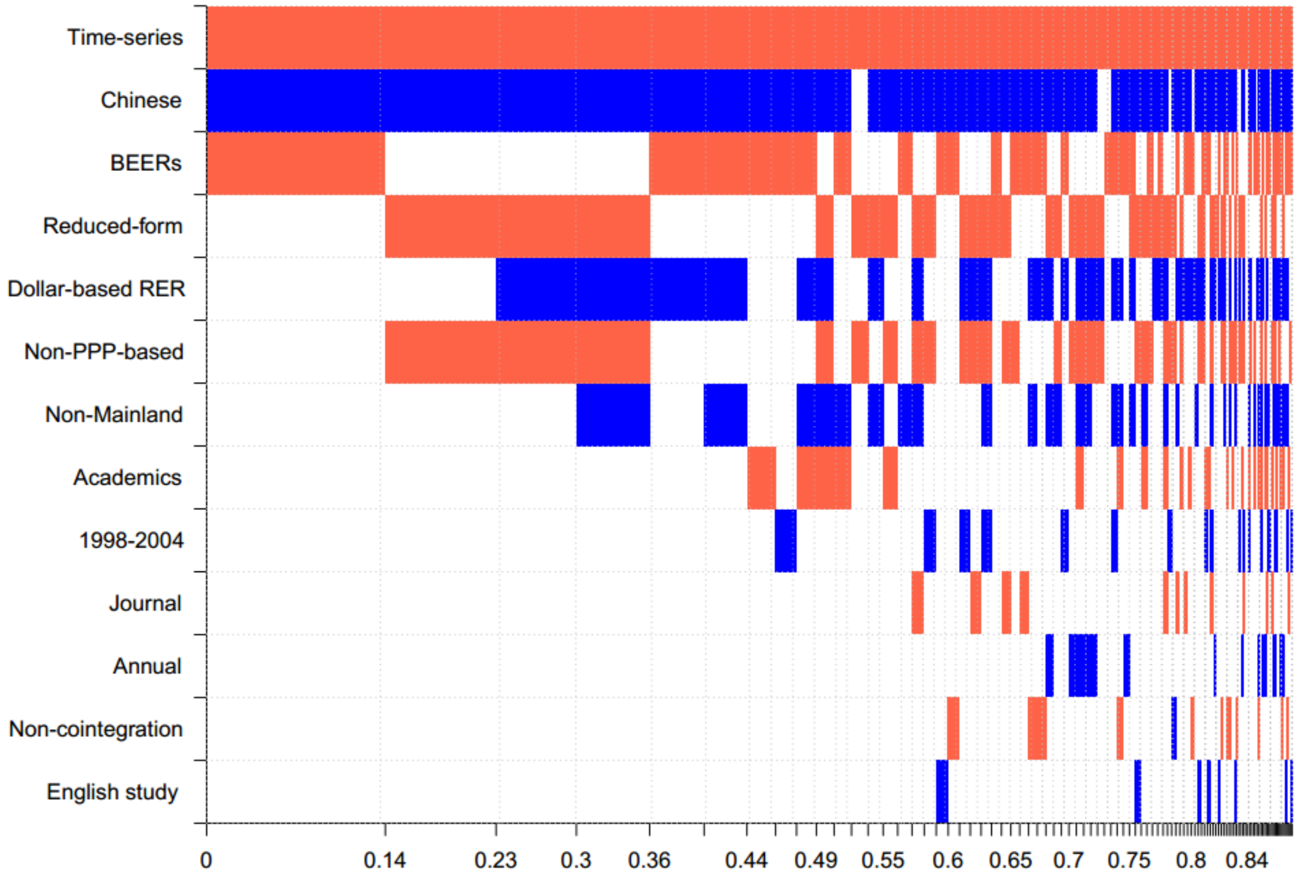


Figure 9. The top 100 models that have the highest posterior model probabilities (based on 13 study characteristic types)



Note: See the Note to Figure 7. The 13 study characteristic types (moderator variables) are listed on the vertical axis in descending orders of their posterior inclusion probabilities. These model specifications are presented from left to right according to their posterior model probabilities from high to low, and the cumulative posterior model probabilities are listed on the horizontal axis.

Appendix

A.1. The Sample of Studies

ID	Study	Publication type	Language
1	Adi and Du (2015)	Journal	English
2	Aflouk, Jeong, Mazier and Saadaoui (2010)	Journal	English
3	Almas, Grewal, Hvide and Ugurlu (2017)	Journal	English
4	Benassy-Quere and Lahreche-Revil (2008)	Journal	English
5	Benassy-Quere, Bereau and Mignon (2009)	Journal	English
6	Benassy-Quere, Lahreche-Revil and Mignon (2011)	Journal	English
7	Chang (2007)	Journal	English
8	Chang (2008)	Journal	English
9	Chang and Qin (2004)	Journal	English
10	Chen (2009)	Journal	English
11	Cheung, Chinn and Fujii (2007)	Journal	English
12	Cheung, Chinn and Fujii (2009)	Journal	English
13	Cheung, Chinn and Fujii (2010)	Journal	English
14	Cheung, Chinn and Fujii (2017)	Journal	English
15	Chou and Shih (1998)	Journal	English
16	Christoph and Hossfeld (2014)	Journal	English
17	Coudert and Couharde (2007)	Journal	English
18	Cui (2013)	Journal	English
19	Frankel (2006)	Journal	English
20	Funke and Rahn (2005)	Journal	English
21	Gan, Ward, Su and Cohen (2013)	Journal	English
22	Garroway, Hacibedel, Reisen and Turkisch (2012)	Journal	English
23	Giannellis and Koukouritakis (2018)	Journal	English
24	Hall, Kenjegaliev, Swamy and Tavlas (2013)	Journal	English
25	Hu and Chen (2010)	Journal	English

26	Lipman (2011)	Journal	English
27	Lü (2007)	Journal	English
28	Nouira, Plane and Sekkat (2011)	Journal	English
29	Peng, Lee and Gan (2008)	Journal	English
30	Schroder (2013)	Journal	English
31	Yang and Bajeux-Besnainou (2006)	Journal	English
32	Yi (2010)	Journal	English
33	You and Sarantis (2011)	Journal	English
34	You and Sarantis (2012a)	Journal	English
35	You and Sarantis (2012)	Journal	English
36	Zhang and Chen (2014)	Journal	English
37	Chen, Deng and Kemme (2008)	Working paper	English
38	Cline (2007)	Working paper	English
39	Cline (2008)	Working paper	English
40	Garton and Chang (2005)	Working paper	English
41	Jeong and Mazier (2003)	Working paper	English
42	Jeong, Bao and Mazier (2007)	Working paper	English
43	Li (2009)	Working paper	English
44	MacDonald and Dias (2007)	Working paper	English
45	Sinnakkannu and Vnair (2010)	Working paper	English
46	Zhang (2010)	Working paper	English
47	Zhang (2012b)	Working paper	English
48	Li (2015)	Master Thesis	English
49	Benassy-Quere, Duran-Vigneron, Lahreche-Revil and Mignon (2004)	Book Chapter	English
50	Cheung, Chinn and Fujii (2012)	Book Chapter	English
51	Cline (2013a)	IIE	English
52	Cline (2013b)	IIE	English
53	Cline (2014a)	IIE	English
54	Cline (2014b)	IIE	English
55	Cline and Williamson (2008)	IIE	English

56	Cline and Williamson (2009)	IIE	English
57	Cline and Williamson (2010a)	IIE	English
58	Cline and Williamson (2010b)	IIE	English
59	Cline and Williamson (2011)	IIE	English
60	Cline and Williamson (2012a)	IIE	English
61	Cline and Williamson (2012b)	IIE	English
62	Subramanian (2010)	IIE	English
63	Shi and Yu (2005)	Journal	Chinese
64	Sun and Sun (2013)	Journal	Chinese
65	Wang (2015)	Journal	Chinese
66	Wang and Cai (2007)	Journal	Chinese
67	Wang and Lin (2013)	Journal	Chinese
68	Wang and Yao (2008)	Journal	Chinese
69	Zhang (2000)	Journal	Chinese

Note: IIE refers to Peterson Institute for International Economics.

A.2. Percentage Misalignment Estimates: Descriptive Statistics

year	Obs	Mean	Std. Dev.	Min	Max
1994	49	25.50	26.5	-4.00	181.00
1995	39	11.50	26.1	-17.40	153.60
1996	39	4.60	28.6	-50.00	145.80
1997	43	8.50	27.7	-30.00	149.30
1998	43	5.20	30.5	-35.00	162.40
1999	54	3.00	30.7	-51.00	173.90
2000	58	11.30	31.6	-36.00	180.20
2001	61	6.70	35.5	-108.00	187.50
2002	64	14.30	28.8	-11.30	194.60
2003	59	21.00	28.9	-6.90	196.00
2004	70	25.00	29.5	-13.30	184.40
2005	87	21.90	27.4	-48.70	164.60
2006	47	20.90	18.7	-5.10	75.80
2007	37	17.40	17.9	-10.00	79.00
2008	52	16.50	19.9	-42.00	78.10
2009	42	9.90	14.4	-15.10	40.20
2010	13	19.70	14.9	-3.90	40.70
2011	51	-3.30	24.1	-87.00	46.00
2012	7	6.70	15.1	-11.40	38.10
2013	8	10.50	10.3	2.00	31.00
2014	14	-1.20	21.5	-27.20	33.30

Note: The table presents, for each year, under the columns “Mean,” “Std. Dev.,” Min,” and “Max” the average, the standard error, the minimum and the maximum of the RMB misalignment estimates, in percentages, of the year listed under the column “year.”

A.3. Study Characteristic types

Study Characteristic types	Description
a) Data Characteristics	
Annual	=1 if annual data are used.
Quarterly	=1 if quarterly data are used.
Monthly	=1 if monthly data are used.
Non-PPP-based	=1 if market based data from, say, IFS, World Bank, or BIS are mainly used.
PPP-based	=1 if PPP-based data derived from ICP surveys are mainly used.
Dollar-based RER	=1 if bilateral real RMB-US dollar exchange rate is used.
REER	=1 if RMB real effective exchange rate is used.
Other RERs	=1 if bilateral real RMB against Japanese yen or euro exchange rate is used.
NER/NEER	=1 if RMB nominal (effective) exchange rate is used.
b) Theoretical and Estimation Specifications	
BEERs	=1 if a model from the family of behavioral equilibrium exchange rate models or the productivity approach is used.
FEERs	=1 if the fundamental equilibrium exchange rate model, IMF macroeconomic balance approach or the nature rate of exchange approach is used.
Penn effect	=1 if the Penn effect approach is used.
PPP	=1 if the absolute or relative PPP framework is used.
Other frameworks	=1 if other frameworks, such as shadow price of foreign exchange approach is used
Time series	=1 if time series technique is used.
Panel	=1 if panel technique is used.
Cross-sectional	=1 if cross-sectional technique is used.
Non-cointegration	=1 if non-cointegration framework is used.
Cointegration	=1 if cointegration framework is used.
Reduced-form	=1 if reduced-form setup is used.

Structural	=1 if structural setup is used.
c) The Publication Attributes	
Journal	=1 if the study is published in a peer-reviewed journal.
Book chapter	=1 if the study is collected from book chapters.
Other types	=1 if the study is neither a journal article nor book chapter
English study	=1 if the study is published in English.
Chinese study	=1 if the study is published in Chinese.
Non-mainland	=1 if no author of the study is affiliated with a mainland China institution.
Mainland	=1 if any one of the authors of the study is affiliated with a mainland China institution.
Chinese	=1 if any one of the authors has a Chinese name and is educated at any level in mainland China.
Non-Chinese	=1 if all authors do not have a Chinese name or not ever educated in mainland China.
Academics	=1 if all authors of the study are affiliated with academic affiliations (e.g. university).
Government	=1 if all authors of the study are affiliated with government affiliations (e.g. central bank) or think tanks (e.g. IIE) or international organizations (e.g. IMF, Asian Development Bank).
Industry	=1 if all authors of the study are affiliated with industry affiliations, such as investment banks and commercial banks.
Mixed	=1 if the authors of the study are affiliated with more than one type of affiliations.
d) Subsample Periods	
1998-2004	=1 when the RMB misalignment estimate falls within the period of 1998 to 2004.
1994-1997	=1 when the RMB misalignment estimate falls within the period of 1994 to 1997.
2005-2008	=1 when the RMB misalignment estimate falls within the period of 2005 to 2008.
2009-2010	=1 when the RMB misalignment estimate falls within the period of 2009 to 2010.
2011-2014	=1 when the RMB misalignment estimate falls within the period of 2011 to 2014.

Note: The BEERs consists of the standard BEER model, the permanent equilibrium exchange rate model (PEER), the equilibrium real exchange rate model (ERER), the Goldman Sachs dynamic equilibrium exchange rate (GSDEER).

A.4. Bayesian Model Averaging

Consider a linear regression model:

$$y = \boldsymbol{X\theta} + \varepsilon\ ;\ \varepsilon \sim N(0, \sigma^2 I_T), \tag{1}$$

where $y = (y_1, \ldots, y_T)'$ is a *Tx1* vector of the dependent variable and ε is a *Tx1* vector of normal random error terms. The *TxK* matrix $\boldsymbol{X}$ = (X_1 . X_2… X_K) contains the *K* potential explanatory variables, and X_j, *j* = 1, .2, …, *K* is a *Tx1* vector of the *j*-th explanatory variable. The coefficients of these *K* variables are in the *Kx1* $\boldsymbol{\theta}$ vector.

Among the potential explanatory variables, which should be used to describe the behavior of *y*? In principle, the *K* potential explanatory variables offer 2^K potential models to consider. Let $\boldsymbol{X}_k$ contains the *k*-th combination of the *K* potential explanatory variables (X_1 . X_2… X_K) and $\boldsymbol{\theta}_k$ is the corresponding coefficient vector. Then, let M_k is the *k*-th of the 2^K models (k = 1, 2, …, 2^K), which is represented by $y = \boldsymbol{X}_k\boldsymbol{\theta}_k + \varepsilon$.

Without a strong (certain) prior of the correct model specification, the selection of an appropriate model to describe the behavior of y involves the model selection uncertainty. Bayesian Model Averaging (BMA) is a way alleviating the problem of model selection uncertainty. The basic idea of BMA is to consider all possible model specifications and make inferences based on a weighted average of posterior probabilities of these models. For model M_k in the model space, its posterior model probability, $p(M_k \mid y, \boldsymbol{X}_k)$, is given by the Bayes' theorem:

$$p(M_k \mid y, \boldsymbol{X}_k) = \frac{p(y \mid M_k, \boldsymbol{X}_k)\, p(M_k)}{\sum_{j=1}^{2^K} p(y \mid M_j, \boldsymbol{X}_j)\, p(M_j)}, \tag{2}$$

where $p(y \mid M_k, X_k)$ is the marginal likelihood of the model M_k, and $p(M_k)$ is the prior model probability. The posterior model probability $p(M_k \mid y, \boldsymbol{X}_k)$ represents how well a model fits the data, and it is analogous to the R^2 estimate or information criteria in frequentist statistics.

The full posterior probability of a coefficient θ_j ; *j* = 1, …., K, vector $\boldsymbol{\theta}$, is

$$p(\theta_j \mid y, \boldsymbol{X}) = \sum_{k=1}^{2^K} p(M_k \mid y, \boldsymbol{X}_k)\, p(\theta_j \mid y, M_k, \boldsymbol{X}_k), \tag{3}$$

which is sum of posterior probabilities of θ_j ($p(\theta_j \mid y, M_k, \boldsymbol{X}_k)$'s) weighted by the respective posterior model probabilities in the model space.

The notion of posterior inclusion probability is used to infer if a coefficient (and the corresponding explanatory variable) should be included in the chosen model. For a given variable X_j (with coefficient θ_j), its posterior inclusion probability is given by

$$PIP_j = \sum\nolimits_{\theta_j \in \boldsymbol{\theta}_k} p(M_k \mid y, \boldsymbol{X}_k)\ ; \tag{4}$$

that is, the sum of posterior probabilities of models that include the variable X_j. The posterior inclusion probability is a measure to assess the (relative) level that the data favor the inclusion of variable X_j in the chosen model. If the posterior inclusion probability of a variable lies between 0.5-0.75, 0.75-0.95, 0.95-0.99 and 0.99-1, then the variable is considered as an acceptable, substantial, strong and decisive effect (Kass and Raftery, 1995; Havranek *et al*., 2015). A variable with posterior inclusion probability smaller than 0.5 is considered ignorable.

The priors on models and priors on coefficients are required to estimate posterior distributions. It is common to employ conservative priors to reflect the situation that the researchers know little about the unknown parameters. We assign a uniform model prior (prior on models) and the unit information prior on Zellner's g-prior (prior on parameters) following Zeugner and Feldkircher (2015), which are quite conservative and reflect unknown true model size and parameter signs.[20]

It is usually neither inefficient nor feasible to compute all potential models since enumerating all models will become time intensive, especially with large number of variables. In our case, we have 26 explanatory variables in basic case and, thus, 2^{26} potential model specifications. The BMS (Bayesian model sampling) package in R provides a Markov Chain Monte Carlo method called Metropolis-Hasting algorithm, which can go through the most important models with high posterior model probabilities (Zeugner and Feldkircher, 2015).

The Metropolis-Hastings algorithm works as follows.[21] For instance, the algorithm first considers a model M_i and calculates its posterior model probability $p(M_i \mid y, \boldsymbol{X}_i)$. Then it draws another model M_i and obtains its posterior model probability $p(M_j \mid y, \boldsymbol{X}_j)$. The algorithm will choose M_j over M_i with the probability $p_{i,j} = \min(1, p(M_j \mid y, \boldsymbol{X}_j) / p(M_i \mid y, \boldsymbol{X}_i))$. If M_j is not selected, the algorithm moves to the next step and draws another model against M_i. If M_j is selected, it replace M_i and the process continues. The distribution of posterior model probabilities will converge if the number of repeated steps is sufficient. For the current study, all BMA computations we use 1,000, 000 burn-ins and 2,000, 000 iterations to ensure a good degree of convergence.

[20] Note that there are alternative settings for priors on models and priors on parameters, such as beta-binomial model prior and benchmark prior on Zellner's g-prior. We also try these settings, and get quite similar results comparing with setting of uniform model prior and the unit information prior on Zellner's g-prior.

[21] See, for example, Zeugner and Feldkircher (2015).