BOFIT Discussion Papers
3 • 2019

Yin-Wong Cheung and Shi He

Truths and myths about RMB misalignment: A meta-analysis

SUOMEN PANKKI 18 11 FINLANDS BANK EUROSYSTEM
BOFIT THE BANK OF FINLAND INSTITUTE FOR ECONOMIES IN TRANSITION

BOFIT Discussion Papers
Editor-in-Chief Zuzana Fungáčová

BOFIT Discussion Papers 3/2019
22.2.2019

Yin-Wong Cheung and Shi He: Truths and myths about RMB misalignment: A meta-analysis

ISBN 978-952-323-265-5, online
ISSN 1456-5889, online

Suomen Pankki
Helsinki 2019

Contents

Yin-Wong Cheung and Shi He

Truths and myths about RMB misalignment: A meta-analysis

Abstract

We conduct a meta-regression analysis of 69 studies that generated 937 renminbi (RMB) misalignment estimates. The Bayesian Model Averaging (BMA) approach is adopted to allow for model selection and sampling uncertainties in assessing effects of study characteristics on these RMB misalignment estimates. Misalignment estimates are found to be influenced by the eight selected study characteristic types in our median probability model. 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 (BMA), clustering effects, median probability model specification, RMB undervaluation, study characteristics.

Yin-Wong Cheung, orcid.org/0000-0002-7564-7402. 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, orcid.org/0000-0001-9986-1799. 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

Acknowledgments
We thank JF Carpantier, Menzie Chinn, David-Jan Jansen, Iikka Korhonen, Juilus Loermann, Błażej Mazur, Shang-Jin Wei, Frank Westermann, James Yetman, and participants at the 12th MIFN International Workshop and the 2018 Conference on “China’s Economic Reforms: Where Do We Stand” for their comments and suggestions. Special thanks to Iikka Korhonen and Shaofang Yu for sharing their data. Finally, we gratefully acknowledge the support of the Hung Hing Ying and Leung Hau Ling Charitable Foundation (孔慶熒及梁巧玲慈善基金) via the Hung Hing Ying Chair Professorship of International Economics (孔慶熒講座教授(國際經濟).

1 Introduction

The long-running debate over currency misalignment in the global economy arouses strong sentiments. In recent history, China's differences with its trading partners over the valuation of its currency, the renminbi (RMB), have dominated headlines. China's current account surplus burgeoned from 1.3% of GDP in 2001 to 9.9% of GDP in 2007, prompting academics and policymakers to scrutinize China's foreign exchange policy. A general view emerged that the RMB was deliberately undervalued and that policy manipulation was bolstering China's trade surpluses (Bergsten and Gagnon, 2012). The US current account deficits and global imbalances were also attributed to RMB undervaluation (Benassy-Quere and Lahreche-Revil, 2008; Cline and Williamson, 2011; Morrison and Labonte, 2013).

As China's current account surpluses mounted, China encountered increasing pressure from the international community – especially the US – to loosen its grip on the RMB and let it appreciate. China's policy response in 2005 allowing the RMB to appreciate did little to calm tensions, however. The US Department of the Treasury (2006), for instance, asserted that "the increased (RMB) flexibility is considerably less than is needed." Over the past two decades, the US has threatened on numerous occasions to label China officially as a currency manipulator, and thereby lay the ground for policy reprisals. Despite its assertion that "China has a long history of pursuing a variety of economic and regulatory policies that lead to a competitive advantage in international trade, including through facilitating the undervaluation of the RMB," nevertheless, the US has not followed through on its currency manipulator threat.[1]

RMB valuation not only remains a contentious issue in the global community (tensions made explicit in the US-China trade war launched in 2018), China's exchange rate policy has implications for its own economy. An artificially depressed currency can overheat the economy and impair monetary policy effectiveness. The crucial first step for all parties concerned is to establish that the RMB is indeed misaligned.

There is a substantial body of studies seeking to estimate the degree of RMB misalignment. After an initial phase dominated by undervaluation estimates, reported empirical estimates of RMB misalignment have spanned a wide over- and undervaluation range. The large disparity in misalignment estimates, not surprisingly, is similar to empirical estimates reported for other economic issues.

The variance in misalignment estimates reflects, among other things, model choice and sample period. For instance, Cheung *et al.* (2007) point out the potential difficulty of pinning down

[1] See United States Department of the Treasury (2018, p. 3). The report states that the US closely tracks the currency practices of countries on its Monitoring List, which includes China, Japan, Korea, India, Germany, and Switzerland. The last time China was cited for manipulating its currency regularly was the period between May 1992 and July 1994 (United States Department of the Treasury, 1992, 1994).

the magnitude of misalignment due to the absence of a consensual equilibrium exchange rate model.[2] Dunaway *et al.* (2009) report that the RMB misalignment estimate is quite sensitive to the assumptions underlying the estimation methods. Based on the available empirical studies, Bineau (2010) and Korhonen and Ritola (2011) found that the reported RMB misalignment estimates were associated with certain study characteristics such as the definition of the RMB exchange rate, the choice of theoretical frameworks and estimation methods, publication types and author affiliations.

That said, the current is not focused on the choice of the equilibrium exchange rate model or the appropriateness of an empirical measure of currency misalignment. Instead, we consider the *variation* of RMB misalignment estimates in terms of the association between their heterogeneity and characteristics of corresponding studies. In this spirit, we extend the meta-analyses of Bineau (2010) and Korhonen and Ritola (2011) to investigate whether the empirical RMB misalignment estimates vary systematically with one or more study characteristics.

The current study extends the existing literature in two ways. First, to the extent possible, our meta-analysis includes empirical studies that report RMB misalignment estimates and their corresponding study characteristics. Our sample includes 69 studies that give 937 RMB misalignment estimates. This wide coverage of studies is done to minimize selection bias.[3] We consider 13 study characteristics that might explain study-to-study heterogeneity of misalignment estimates. While comparable to those considered by e.g. Bineau (2010), Egert and Halpern (2006), and Korhonen and Ritola (2011), these study characteristics offer fairly 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 our case) and the attributes of empirical studies, there is no strong theory on what key study characteristic types are appropriate. Indeed, 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. BMA is an established procedure that explicitly accounts for uncertainties due to model selection and estimation procedure in generating inference on the parameter estimates of interest.[4]

In anticipation of results, our BMA-based meta-analysis shows that the reported RMB misalignment estimates are affected by certain study characteristics. For instance, study characteristic

[2] The seminal article of Meese and Rogoff (1983) documents the inability of economic models to explain exchange rate movement – a finding echoed by e.g. Cheung *et al*. (2005, 2018), and Rossi (2013).
[3] Bineau (2010) and Korhonen and Ritola (2011) cover, respectively, 17 and 30 studies, and 130 and 99 misalignment estimates in their exercises.
[4] The BMA approach is adopted to account for modeling uncertainty in the meta-analysis by, for example, Fernandez *et al.* (2001) on cross-country growth regressions and Irsova and Havranek (2013) on the determinants of horizontal spillovers from foreign direct investment.

types that include the use of PPP-based data, a cross-sectional setting, and a structural setup tend to yield strong evidence of RMB undervaluation. Studies that use data on real effective exchange rate or nominal (effective) exchange rate, studies prepared by a group of authors from academics, government, or industry, and estimates for the period of 2009 to 2014 tend to find weak evidence of RMB undervaluation. Moreover, the use of the least squares regression technique (with and without controlling for clustering on study) yields different sets of significant study characteristic types. Finally, the significance of study characteristics is 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 estimates and graphical analyses. Section 3 reports the results from the BMA-based meta-analysis. Section 4 provides additional results obtained from adopting the least squares approach with and without controlling for clustering on study, as well as results from an alternative choice of benchmark study characteristics. Section 5 offers concluding remarks.

2 Preliminaries

After China launched its reform initiative in 1978, it has experienced phenomenally strong economic growth and rapidly integrated with the global economy. This historically unprecedented progress has had substantial implications for the Chinese economy and the world. China's emergence onto the global stage also put its RMB policy in the limelight, triggering considerable interest in estimating the equilibrium value and the misalignment of the Chinese currency. Chou and Shih (1998) present an early empirical study on RMB misalignment. The debate on the RMB valuation gathered momentum as China's trade surplus surpassed 4% of GDP in 2005 and reached 9.9% in 2007. In recent years, there is a slowdown of the study of RMB misalignment; there are only three journal articles addressed on RMB misalignment in 2017 and 2018 (i.e., Almas *et al*., 2017; Cheung *et al.* 2017; and Giannellis and Koukouritakis, 2018) in our study.

In the last two decades, researchers produced a plethora of empirical studies on RMB misalignment based on various models and estimation methods, and covering different time periods. These empirical studies generated conflicting inferences about the level of RMB misalignment, with misalignment estimates ranging from substantial undervaluation to significant overvaluation. We adopt the BMA-based meta-analysis approach to study the implications of study characteristics for the observed heterogeneity of study-to-study RMB misalignment estimates.

2.1 Sample of studies

Where possible, we include studies that report quantitative inferences about the RMB misalignment in order to alleviate selection bias. Our raw sample of studies was constructed as follows. For studies in English, we include studies considered by Bineau (2010) and Korhonen and Ritola (2011). Next, we search 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 cite any of these five studies.[5] For studies in Chinese, we sort through the top ten Chinese economics journals using the China National Knowledge Infrastructure (CNKI) search engine.[6] Finally, we examine studies published between 2015 and 2018 obtained from the previous steps and look for relevant references.[7] This effort gives us a total of 283 studies in our preliminary sample.

In the preliminary sample, looking for studies that report quantitative estimates of RMB misalignment on or after 1994, we identify a total of 69 studies of which 62 are English publications and 7 Chinese publications. These studies constitute the sample of studies examined in the following analyses. They are listed in Appendix A1.

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.[8] The terms “RMB misalignment estimates” or “misalignment estimates” here refer to these quantitative estimates of RMB misalignment.

Second, we focus on misalignment estimates on or after 1994. China instituted a major change of its exchange rate policy in January 1994, replacing the dual-exchange rate arrangement to a managed exchange rate with the US dollar. 2014 is the last year in our sample for which we have RMB misalignment estimates.[9]

---

[5] The five identified studies are Chang and Qin (2004), Cheung et al. (2007), Chou and Shih (1998), Funke and Rahn (2005), and Zhang (2001). These studies have garnered, respectively, 122, 288, 158, 224, and 224 citations (Google Scholar searches as of June 2018).

[6] The top ten (and most influential) Chinese academic journals on economics 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,” “数量经济技术经济研究,” “中国农村经济,” and “经济学动态.”

[7] The articles published during 2015–2018 are Agya and Jun (2015), Almas et al. (2017), Cheung et al. (2017), Giannellis and Koukouritakis (2018), Li (2015), and Wang (2015).

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

[9] Bineau (2010) and Korhonen and Ritola (2011) include misalignment estimates from 1975 to 2008 and from 1998 to 2009.

Figure 1 displays the frequencies of the selected studies according to the years that they were published. There are only a handful of studies before the turn of the millennium. Chou and Shih (1998) and Zhang (2000) are the two journal articles published between 1998 and 2000. The bulk of selected studies are published between 2007 and 2013, i.e. after China's trade surplus soared. The latest studies in our sample are Almas *et al.* (2017), Cheung *et al.* (2017), and Giannellis and Koukouritakis (2018).

The box plots of RMB misalignment estimates in individual years are presented in Figure 2. For easy reference, the legend of the box plot (a.k.a. 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 trends upward to reach its peak value in 2004 before drifting lower. After a brief pick-up in 2010, it resumes 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 a time-varying pattern: 2004 has the largest inter-quantile range estimate, while 2001 has the largest range between extreme undervaluation and overvaluation estimates. With the exceptions of 2009, 2010, 2013, and 2014, there are extreme misalignment estimates reported for individual years in the sample. Extreme undervaluation estimates outnumber extreme overvaluation estimates. In general, the range of year-to-year variability increases in the early part of the sample, then diminishes over time.[10]

2.2 Study characteristics

Meta-analysis is used here to investigate the potential roles of features of studies that might explain variations in RMB misalignment estimates. To this end, we collect information on 13 study characteristics and group them into four categories: 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 Appendix A3.[11]

2.2.1 Data characteristics

This category comprises three study characteristics. Specifically, we coded the studies according to whether annual, quarterly, or monthly data are used; whether the data are mainly derived from PPP-based data such as International Comparison Program (ICP) surveys; and whether the RMB real

[10] Appendix A.2 presents, for each year, the mean, standard error, minimum, and maximum of misalignment estimates.
[11] 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.

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

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

Compared to studies that use 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 stronger evidence of undervaluation (Figure 3b).

Misalignment estimates of the RMB-US dollar real exchange rate and 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 (Figure 3c).

2.2.2 Theoretical and estimation specifications

There are four study characteristics in this category. For the study characteristic of theoretical settings, we consider five types: "BEERs," "FEERs," "Penn effect," absolute or relative PPP framework ("PPP"), and "Other frameworks."[12] For the three estimation specification characteristics, we classify the studies according to whether a panel, cross-sectional, or time series approach is adopted, whether a cointegration framework is used, and 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 4a). The Penn effect regression generates the next largest number of misalignment estimates. The Penn-effect-based estimates relative to other methods yield the largest median of undervaluation estimates and the largest range of estimates (as evidenced 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 the time series data (Figure 4b). While studies using time series

[12] 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.

data yield similar shares of overvaluation and undervaluation extreme estimates, those using cross-sectional data mostly yield undervaluation extreme estimates.

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

2.2.3 Publication attributes

A study characteristic in this category is the venue of publication. We consider whether the study has been published as an academic journal article, a book chapter, or another format such as working paper. Another characteristic considered is whether the study has been published in English or Chinese language.

Figures 5.a and 5.b show that most of the misalignment estimates in our sample come from academic journal articles and 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, but give the largest median estimate in this study characteristic.

The other three study characteristics in this category relate to the author(s) of the selected studies. We consider whether any of the authors has an affiliation with an institution in mainland China, whether any of the authors has a Chinese name (and has been educated in China at any education level),[13] and whether all of the authors of a particular study only have academic affiliations, "Government" affiliations,[14] or industry affiliations (e.g. investment bank and commercial bank), rather than a mix of such affiliations.

Figures 5c and 5d show that studies with 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 relatively large proportion of extreme misalignment estimates are found among studies with authors affiliated with a mainland China institution or Chinese.

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

[13] We checked their on-line curriculum vitae for their education histories.
[14] "Government" refers to government affiliations (e.g. central banks), think tanks (e.g. Peterson Institute for International Economics), or international organizations (e.g. IMF and Asian Development Bank).

2.2.4 Subsample periods

China modified its exchange rate policy several times during our sample period. It replaced the dual-exchange rate arrangement with a policy targeting the US dollar in January 1994. It then moved to a *de facto* dollar peg after the 1997 Asian Financial crisis. The "managed floating exchange rate regime" adopted in July 2005 was abandoned for a stable RMB/dollar rate policy in the midst of the Global financial crisis in 2008, only to have the "managed floating exchange rate regime" reestablished in 2011. As a pegged rate policy arguably hinders the exchange rate adjustment process, it has 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.a 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 for the period 2011–2014 is one – the smallest of these five subsample periods.

Figure 6b displays the dollar-based RMB exchange rate and averaged RMB misalignment estimate for each year. The RMB exchange rate is quite stable until it enters a steady appreciation trend in 2005. The yearly average of RMB misalignment estimate, after an initial drop, increases and reaches its peak (25.01%) in 2004. Thereafter, the average shows a declining trend. Apparently, the 2005 reform has some implications for these RMB undervaluation estimates.

The box plots offer some circumstantial evidence on implications of study characteristics for RMB misalignment estimates. The observed differential effects across characteristic types of a given study characteristic, however, can be influenced by the interaction 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 suggest RMB misalignment estimates are associated with certain study characteristics. An astute reader, of course, requires additional statistical evidence to confirm the roles of such study characteristics. To develop this evidence, we adopt the regression framework:

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

to examine the study-to-study variation of RMB misalignment estimates. 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}, 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 ("Annual," "Quarterly," and "Monthly") leads to perfect collinearity in the presence of a regression intercept term (or other qualitative response variables). Thus, we define for a given study characteristic a "benchmark" characteristic type as a reference point for assessing the effects of study characteristics. In this section, we identify the benchmark type for each study characteristic as it is the study characteristic type adopted by most studies. The 13 selected benchmark characteristic types are bolded in Appendix A3.[15]

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, so which set of variables for the sake of practicality should be included in the empirical analysis? Despite some of the anecdotal evidence from the figures in the previous section, we arguably lack a robust theory on how to select 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 BMA approach that 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. Based on a parameter's posterior distribution, we obtain its posterior mean and posterior standard error.[16] See Appendix A4 for a discussion of the BMA methodology.

[15] In the next section, an alternative set of benchmark types is considered.
[16] As a heuristic comparison, the PIP analogizes the p-value, posterior mean the point estimate, and posterior standard error the standard error under the frequentist approach.

3.1 Basic BMA results

To assess which 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: the uniform prior probability on the model space (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.[17]

Figure 7 presents information of the top 6,000 model specifications with the highest posterior model probabilities.[18] The 26 study characteristic types are listed on the vertical axis in descending order of their PIPs. 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 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. These 6,000 models account for about 90% of the probability on the model space.

Two study characteristic types, “2011-2014” and “Cross-sectional,” are included in these top 6,000 model specifications. The “2011-2014” characteristic type displays a consistently negative sign (red cell; lighter color in grayscale), while the “Cross-sectional” characteristic shows a consistently positive sign (blue cell; darker color in grayscale) in these specifications.

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. This indicates that these study characteristic types can be correlated, and thus can display effects in a multivariate framework different from the descriptive statistics depicted in box plots.

[17] The extreme values visualized in the box plots in the previous section are not excluded from the exercise. Instead, their relevance is determined via posterior probabilities based on the priors and the likelihood function.

[18] 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).

Table 1 presents statistics derived from its full posterior probability of the parameter and the posterior model probability for each study characteristic type. Under the column label “PIP,” we report the PIP 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, or decisive effect if its PIP falls between 0.5–0.75, 0.75–0.95, 0.95–0.99, and 0.99–1, respectively. If the PIP is less than 0.5, the study characteristic type is considered ignorable. In Table 1, bolded figures indicate that the study characteristic type has an estimated PIP greater 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 (i.e. the study characteristic type tends to yield a high level of RMB undervaluation), and a value of zero implies the parameter is negative (that is, the study characteristic type tends to yield a low level of RMB undervaluation). 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 characteristic types have PIP values noticeably below the 0.5 threshold, and posterior means that are close to zero and small compared with their corresponding posterior standard errors. 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. This finding is in line with Bineau (2010) and Korhonen and Ritola (2011).

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 PIP 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, i.e. 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. China’s productivity growth underlying its spectacular economic performance is not properly reflected in market exchange rates and prices. To the extent that relative productivity growth has implications for purchasing power, and hence the 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 other types of exchange rate such as nominal (effective) exchange rates tend to yield weak evidence of RMB undervaluation. These two exchange rate characteristic types have a value of zero under "Sign" and have PIP values above 0.5, suggesting 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 characteristic types "FEERs," "Penn effect," "PPP," and "Other frameworks" have PIP values less than 0.5 and a very small posterior mean relative to the posterior standard error. Thus, these theoretical frameworks are unlikely 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 unlikely to be the source of study-to-study variations of misalignment estimates as the "Cointegration" type in Table 1 has a PIP value of just 0.108.

Under the category of theoretical and estimation specifications, "Cross-sectional" and "Structural" are the two study characteristic types with PIP values greater than 0.9 and tend to generate RMB undervaluation estimates. Under a time-series model specification, the (estimated) equilibrium exchange rate is typically given by the average over time (conditional on regressors), so overvaluation and undervaluation estimates are almost invariably reported. When cross-sectional data are used, it is implicitly assumed (conditional on regressors) that 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, say, is undervalued for an extended period and not at the (estimated) equilibrium value. The significance of the "Cross-sectional" characteristic type suggests the estimated equilibrium exchange rate based on cross-country averages exceeds the RMB rate.

Our significant "Structural" characteristic type result is comparable to Wang and Yao (2008), who find that the structural setup usually gives larger misalignment estimates than the reduced-form approach.

The “Mixed” characteristic type under the category of publication attributes is the only characteristic type that has a PIP estimate greater 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 2009–2010 and 2011– 2014 tend to be different from those for other periods. Both cases have a value of zero under the “Sign” column, with the 2011–2014 misalignment estimates are more likely to display a small undervaluation value. The PIP values, however, indicate that estimates generated for assessing RMB misalignment in 2011–2014 should have a higher chance to affect misalignment estimates than those for the 2009–2010 period. This finding accords with the anecdotal evidence of a strengthening RMB and narrowing Chinese current account surplus observed after the global financial crisis (Yue *et al.*, 2016).

Our meta-analysis based on the BMA method offers evidence that the reported RMB misalignment estimates are associated with certain study characteristics, including the property of the data used, the choice of theoretical and estimation methods, author’s affiliations, and the periods for which the estimates are generated. The empirical effects of the study characteristics have been accounted for in both sampling variations and uncertainties related to 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 PIP larger than 0.5. These eight variables constitute our median probability model.[19]

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 specification? 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, indicate that the median probability model yields good predictions.

Adopting the median probability model specification, we obtain the density plot of the RMB misalignment estimate depicted in Figure 8. The 2.5%, 50%, and 97.5% quantiles of the estimated density are, respectively, -0.342, 0.135 and 0.611. Thus, while the median of the estimate is

[19] 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. They include the same set of eight study characteristic types (Figure 7).

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, i.e. a study that possesses these eight study characteristic types.[20] The resulting 2.5%, 50%, and 97.5% quantiles of the RMB misalignment estimate are, respectively, -0.787, -0.267, and 0.261. The change of the assumed values of study characteristic types decreases the median of the RMB misalignment estimate from positive to negative, so 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 specification into two groups. The first comprises three study characteristic types with a positive sign. The second has 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 -0.090, 0.426, and 0.943. For the case of five study characteristic types with a negative sign, the 2.5%, 50%, and 97.5% quantiles are -1.204, -0.678, and -0.153. Thus, 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. If the hypothetical study uses the five identified "negative" study characteristic types, there is a better than 95% chance that the RMB is found to be overvalued.

An average model approach that considers all the study characteristic types is an alternative way to generate the prediction of RMB misalignment (Eklund and Karlsson, 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 one that includes the least commonly used study characteristic types, and studies that 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.[21]

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 equipped with the five identified negative characteristic

[20] 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.

[21] Results of these cases are available from the authors. See footnote 20.

types, it is hard to draw a definitive inference to reject the hypothesis of the RMB is not misaligned after controlling for model selection and estimation uncertainties. Notably, Cheung *et al.* (2007), Dunaway *et al.* (2009), and Schnatz (2011) all 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 obtain additional insight on the effects of study characteristics on the reported RMB misalignment estimates, we present results from the least squares approach and an alternative benchmark specification.

4.1 Regression analysis

In this subsection, we present results based on our 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 yield a statistically significant coefficient estimate with a sign consistent with the one given in Table 1. The magnitudes of these least squares estimates, however, can be 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 PIP 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,” which all have PIP values above 90%. Recall that under the BMA approach, the PIP 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 that 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 that the least squares approach overstates 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 heteroskedasticity-robust standard errors reported in columns (1) and (2) can overstate the level of 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, the number of significant study characteristic types of the former model drops to 7 from 19 while the latter drops to 7 from 8. Thus, 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. Only three significant study characteristic types ("BEERs," NER/NEER," "2011-2014"), however, are common to both 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 happens if different benchmarks, e.g. the quarterly and monthly data instead of annual data are the designated benchmark characteristic type of the data frequency study characteristic? In this sub-section, we reassess the results after inter-changing the roles of benchmark and non-benchmark characteristic types used in the previous section. In other words, 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 (bolded figures 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. Together they 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 weak evidence of RMB undervaluation,[22] and a study that has an author with a Chinese name is likely to generate a strong evidence. In addition to these three types, the characteristic type "Reduced-form" also garners a PIP larger than 0.5 (Table 3). For this group of 13 study characteristic types, the median probability model has four explanatory variables: "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 of the distribution of the resulting RMB misalignment estimate are given by -0.374, 0.135, and 0.643. When the four variables assume the value of one, then 2.5%, 50% and 97.5% quantiles become -0.481, 0.028, and 0.537. 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.[23]

Table 4 presents the least squares regression results. 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 PIP of the four explanatory variables, is not significant. The other three study characteristic types are statistically significant with signs in accordance with those given in Table 3. In column (2), where the average model specification comprises all the 13 characteristic types, we see 12 are significant.

The standard errors that are robust to clustering on study are presented under columns (3) and (4). Controlling for clustering appreciably 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 the findings of 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

---

[22] Similar negative "BEERs" and "Time series" effects are reported in Korhonen and Ritola (2011).

[23] 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).

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 a time series specification typically identifies the equilibrium exchange rate with the average (conditional on regressors) over the sample period. 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, 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.

Note that the adjusted R^2 estimates indicate that specifications in Table 2 possess better explanatory power than their counterparts in Table 4.

Besides these extra regressions, we considered a setup in which some of the 26 characteristic types were combined to increase the number of observations per characteristic type.[24] We also explicitly accounted for outliners in the BMA and least squares regressions.[25] All these alternative attempts yielded results qualitatively similar to those reported. The RMB misalignment estimate displays different study characteristic determinants under different specifications, and is insignificantly different from zero under various hypothetical combinations of study characteristics.

Our exercise also assessed publication bias arising from the predisposition of researchers and reviewers to accept empirical results that are consistent with the conventional view or "statistically significant" (Card and Krueger, 1995; Stanley and Doucouliagos 2012). In addition to direct evidence from results based on different classifications of publication attributes, alternative choices of, say, date frequencies and methodologies can contribute to publication bias. On publication attributes, "Mixed" in Table 1 and "Chinese" in Table 3 are the two instances that indicate the possibility of publication bias, and the venue of publication has no implication. Our findings are differ from Korhonen and Ritola (2011), who find that journal-published papers tend to report high levels of RMB undervaluation. The difference is likely due to differences in coverages of studies and sample periods.

[24] Specifically, the "Quarterly" and "Monthly" characteristic types are combined to form the non-annual type, the "Book chapter" and "Other type" characteristic types are combined to form the non-journal type, and the "Government", "Industry" and "Mixed" characteristic types are combined to form the non-academic type. As a result, the modified setup has 22 study characteristic types.

[25] We considered two ways of defining outliners. Outliners are the misalignment estimates that are either larger than 100% in absolute value, or beyond the lower/upper adjacent values of the box plot of all the 937 observations.

5 Concluding remarks

We adopt the BMA approach to conduct a meta-analysis of the effect of study characteristics on empirical RMB misalignment estimates. An 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 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. For all but one exception of the hypothetical cases considered, we found that the RMB misalignment estimate is insignificantly different from zero. Thus, we cannot reject the notion of the RMB is not misaligned.

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. Pooling under the meta-analysis setup should aggregate information and improve the precision of estimation. 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.

These findings imply that great caution should accompany any assessment and interpretation of reported RMB misalignment estimates. Cheung *et al*. (2007), Dunaway *et al*. (2009), and Schnatz (2011), for example, noted that the evaluation of misalignment is hindered by the absence of a consensual equilibrium exchange rate model, substantial data revisions, and sensitivity to small changes in assumptions underlying empirical specifications. Both the theory and empirical data are not sufficiently informative for deciphering the equilibrium value, and thus 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. For academics and policymakers, prudence should be a crucial element of asserting the exact level of RMB misalignment and recommending related policy remedies.

References

Agya, A.A. and Jun, D. (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, I., Grewal, M., Hvide, M. and Ugurlu, S. (2017). The PPP approach revisited: A study of RMB valuation against the USD. Journal of International Money and Finance 77: 18-38.

Barbieri, M.M. and Berger, J.O. (2004). Optimal predictive model selection. The Annals of Statistics 32(3): 870-897.

Benassy-Quere, A., Lahreche-Revil, A. and Valerie, M. (2008). Is Asia responsible for exchange rate misalignments within the G20? Pacific Economic Review 13(1): 46-61.

Bergsten, C.F. and Gagnon, J.E. (2012). Currency Manipulation, the US Economy, and the Global Economic Order. Peterson Institute for International Economics Policy Brief No. 12-25. Peterson Institute for International Economics: Washington DC.

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

Card, D. and Krueger, A.B. (1995). Time-series minimum-wage studies: A meta-analysis. American Economic Review 85 (2): 238-243.

Chang, G.H. and Qin, S. (2004). How much is the Chinese currency undervalued? A quantitative estimation. China Economic Review 15(3): 366-371.

Cheung, Y-W., Chinn, M.D. and Fujii, E. (2007). The overvaluation of renminbi undervaluation. Journal of International Money and Finance 26(5): 762-785.

Cheung, Y-W., Chinn, M.D. and Fujii, E. (2010). Measuring misalignment: Latest estimates for the Chinese yuan. In: Evenett, S.J. (ed.). The US-Sino Currency Dispute: New Insights from Economics, Politics and Law, Chapter 10. Centre for Economic Policy Research: London.

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

Cheung, Y-W., Chinn, M.D. and Pascual, A.G. (2005). Empirical exchange rate models of the nineties: Are any fit to survive? Journal of International Money and Finance 24(7): 1150-1175.

Cheung, Y-W., Chinn, M.D., Pascual, A.G. and Zhang, Y. (2018). Exchange rate prediction redux: New models, new data, new currencies. Journal of International Money and Finance. (Forthcoming, doi: 10.1016/j.jimonfin.2018.03.010)

Cheung, Y-W. and Fujii, E. (2014). Exchange rate misalignment estimates – Sources of differences. International Journal of Finance and Economics 19(2): 91-121.

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

Clark, P.B. and MacDonald, R. (1999). Exchange rates and economic fundamentals: A methodological comparison of BEERs and FEERs equilibrium exchange rates. In: MacDonald, R. and Stein, J.L. (eds.), Equilibrium Exchange Rates, Chapter 10. Springer: Dordrecht.

Cline, W.R. and Williamson, J. (2010). Estimates of Fundamental Equilibrium Exchange Rates. Peterson Institute for International Economics Policy Brief No. 10-15. Peterson Institute for International Economics: Washington DC.

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

Egert, B. and Halpern, L. (2006). Equilibrium exchange rates in Central and Eastern Europe: A meta-regression analysis. Journal of Banking & Finance 30(5): 1359-1374.

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

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

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

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

Giannellis, N. and Koukouritakis, M. (2018). Currency Misalignments in the BRIICS Countries: Fixed vs. Floating Exchange Rates. Open Economies Review 29(5): 1-29.

Havranek, T., Horvath, R, Irsova, Z. and Rusnak, M. (2015). Cross-country heterogeneity in intertemporal substitution. Journal of International Economics 96(1): 100-118.

Irsova, Z. and Havranek, T. (2013). Determinants of horizontal spillovers from FDI: Evidence from a large meta-analysis. World Development 42: 1-15.

Meese, R. and Rogoff, K. (1983). Do they fit out of sample? Journal of International Economics 14: 3-24.

Morrison, W.M. and Labonte, M. (2013). China's currency policy: An analysis of the economic issues. Congressional Research Service.

Kass, R.E. and Raftery, A.E. (1995). Bayes Factors. Journal of the American Statistical Association 90(430): 773-795.

Korhonen, I. and Ritola, M. (2011). Renminbi Misaligned Results from Meta-Regressions. In: Cheung, Y.W. and Ma, G. (eds.), Asia and China in the Global Economy, Chapter 4. World Scientific Publishing Company: Singapore.

Ley, E. and Steel, F.J. (2012). Mixtures of g-priors for Bayesian model averaging with economic applications. Journal of Econometrics 171(2): 251-266.

Li, C. (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, A, Mazier, J and Saadaoui, J. (2012). Temporal dimension and equilibrium exchange rate: A FEER/BEER comparison. Emerging Markets Review 13(1): 58-77.

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

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

Stanley, T.D. and Doucouliagos, H. (2012). Meta-regression analysis in economics and business. Routledge: London.

Tukey, J.W. (1977). Exploratory data analysis. Addison-Wesley: Boston.

United States Department of the Treasury (1992). Report to the Congress – Report to Congress on International Economic and Exchange Rate Policies, May 1992. U.S. Department of the Treasury, Office of International Affairs.

United States Department of the Treasury (1994). Interim Report to the Congress on International Economic and Exchange Rate Policies, July 1994. U.S. Department of the Treasury, Office of International Affairs.

United States Department of the Treasury (2006). Report to the Congress – Report to Congress on international economic and exchange rate policies, December 2006. U.S. Department of the Treasury, Office of International Affairs.

United States Department of the Treasury (2018). Report to the Congress – Macroeconomic and foreign exchange policies of major trading partners of the United States, October 2018. U.S. Department of the Treasury, Office of International Affairs.

Yue, L.H., Qiang, J. and Kai, T.Y. (2016). Determination of renminbi equilibrium exchange rate, misalignment, and official intervention. Emerging Markets Finance and Trade 52(2): 420-433.

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

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

Wang, B. 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) 4: 1277-1302. Available from CNKI.

Wang, Z. 王泽填 and Yao, Y. 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, X. 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 8: 13-24. Available from CNKI.

Tables and figures

Table 1 BMA results based on 26 study characteristic types

	PIP	Post Mean	Post SD	Sign
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**
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**
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**
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

Note: “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). Bolded figures indicate that the corresponding study characteristic type has an estimated PIP greater than 0.5.

Table 2 OLS results: Robustness check – I

	1	2	3	4
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)
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)
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)

	1	2	3	4
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: Least squares regression results are presented here. Columns 1 and 3 give 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. White heteroskedasticity-robust standard errors are presented in parentheses under the coefficient estimates in columns 1 and 2. Standard errors robust to clustering by studies are presented in parentheses under the coefficient estimates in columns 3 and 4. See discussion in main text for details.

Table 3 BMA results based on 13 study characteristic types

	PIP	Post Mean	Post SD	Sign
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
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**
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
Subsample periods				
1998-2004	0.119	0.004	0.012	1.000
(Intercept)	1.000	0.231	NA	NA

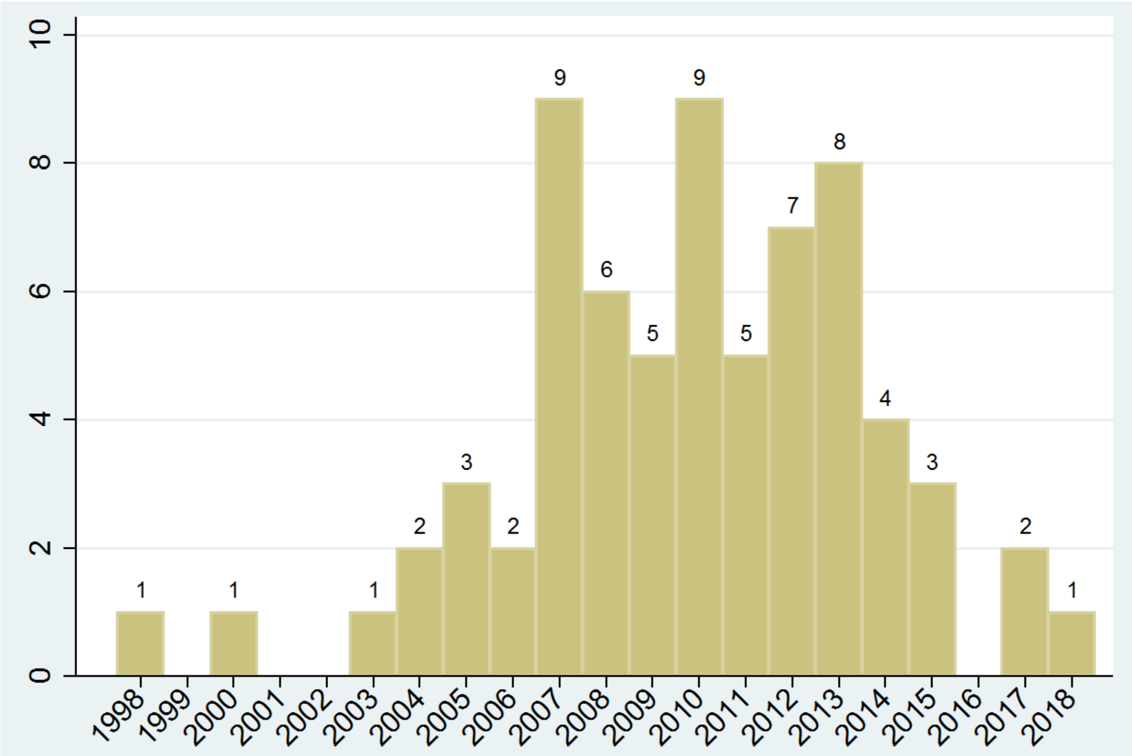
Note: See the note for Table 1.

Table 4 OLS results: Robustness check – II

	1	2	3	4
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)
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)
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)
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

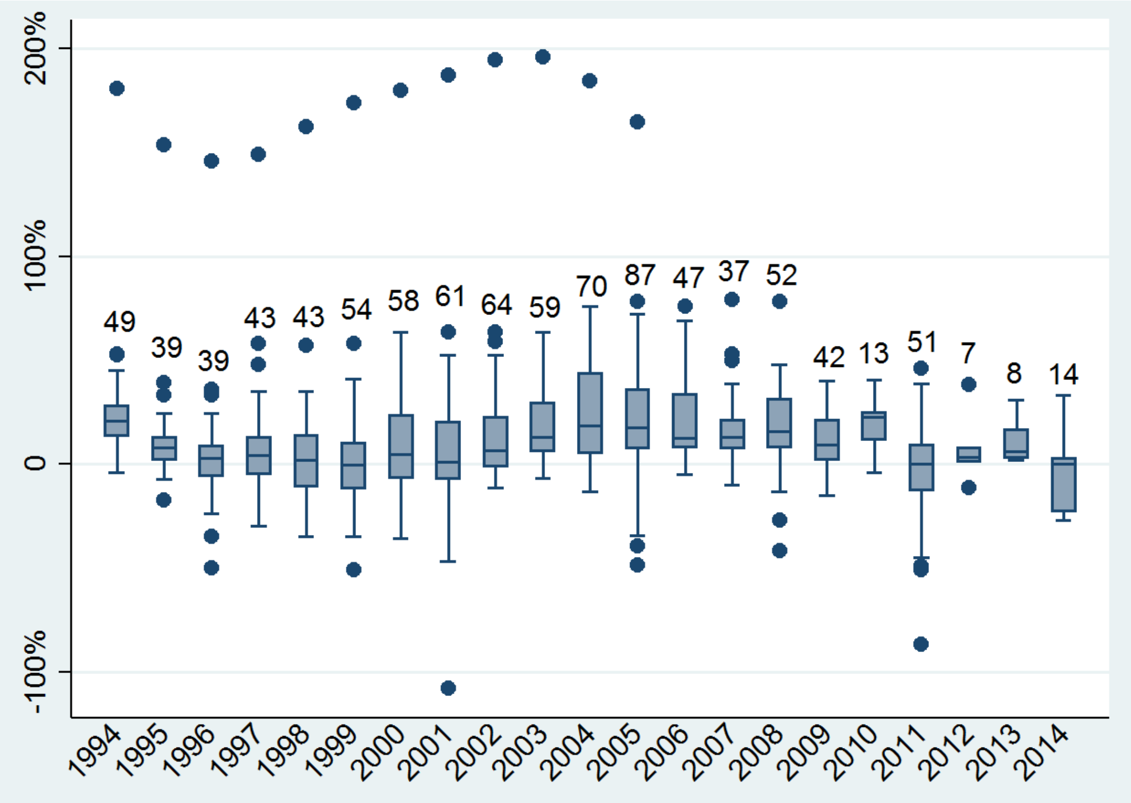
Note: See the note for Table 2.

Figure 1 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 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 follows Tukey (1977):

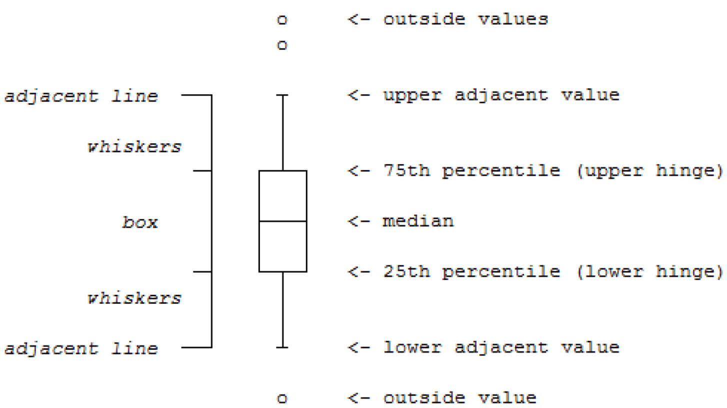
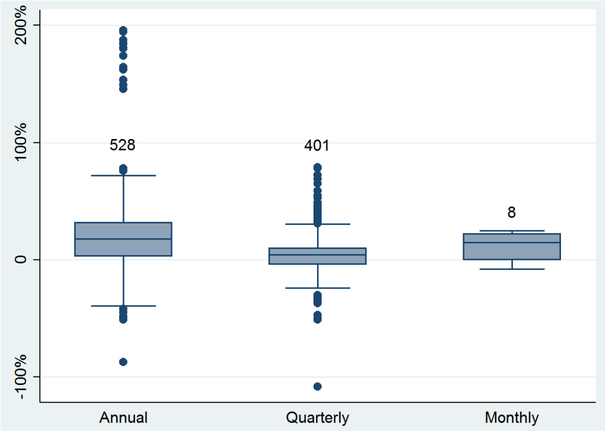
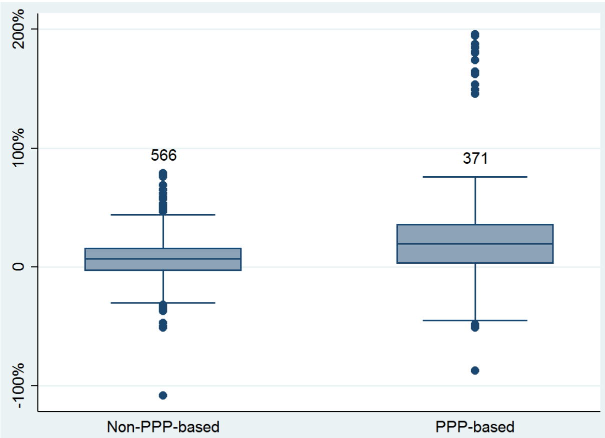


Figure 3 Box plots of RMB misalignment estimates: Data characteristics

3a. Data frequency



3b. Data type



3c. Definition of RMB exchange rate

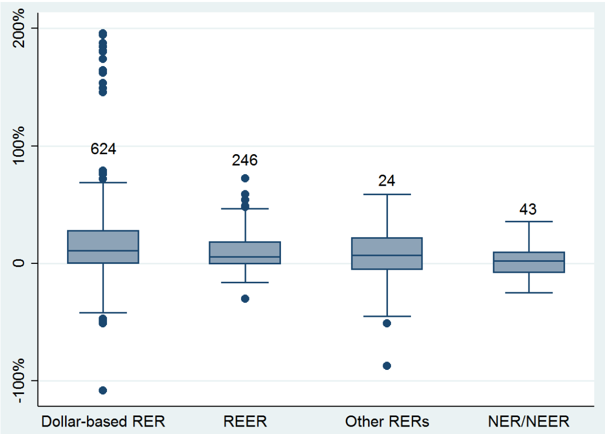
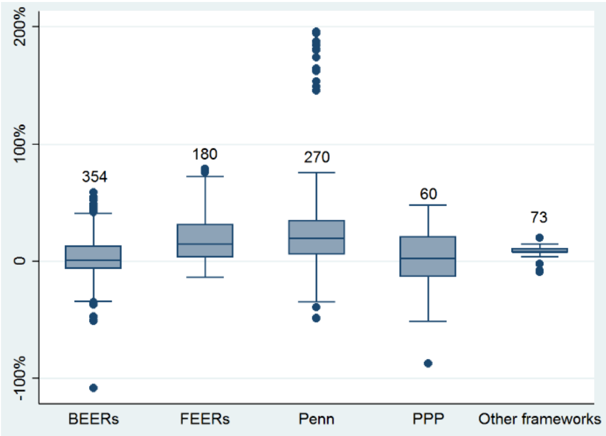
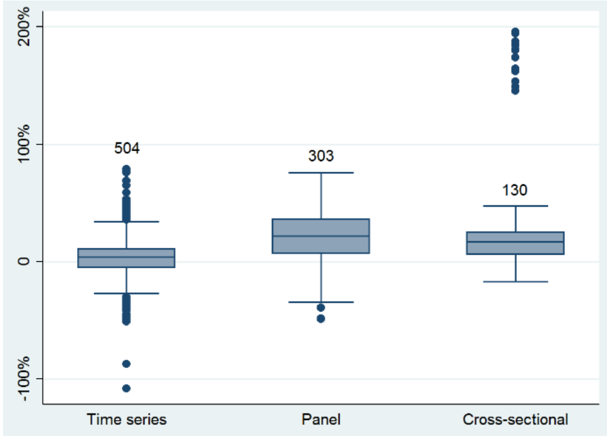


Figure 4 Box plots of RMB misalignment estimates: Theoretical and estimation specifications

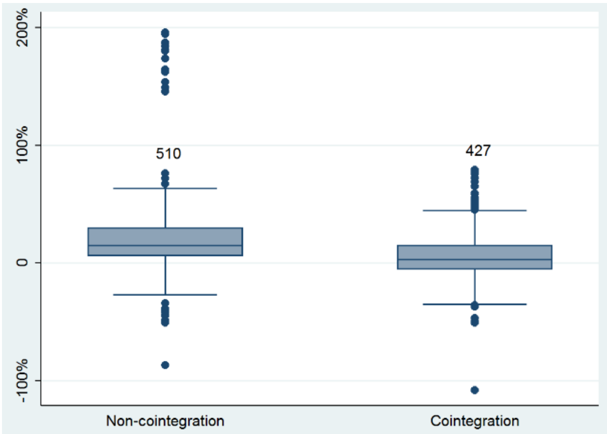
4a. Theoretical specifications



4b. Estimation method (1)



4c. Estimation method (2)



4d. Estimation method (3)

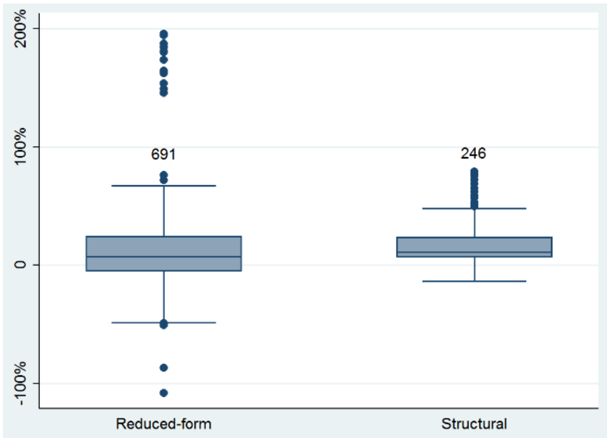
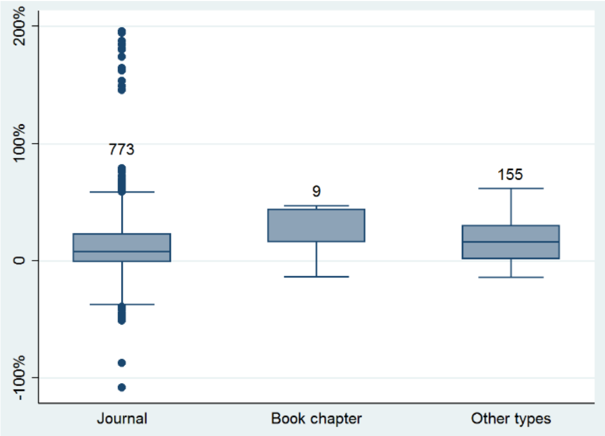
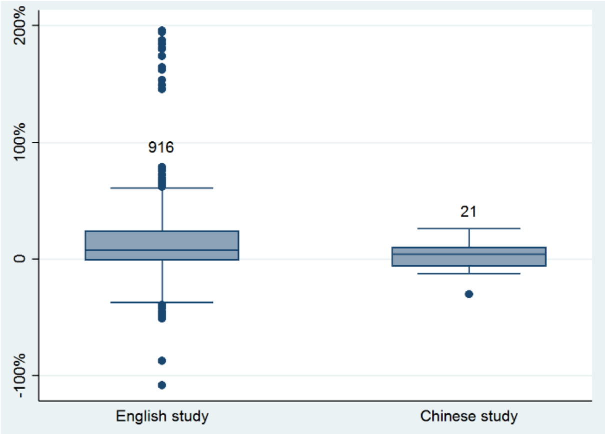


Figure 5 Box plots of RMB misalignment estimates: Publication attributes

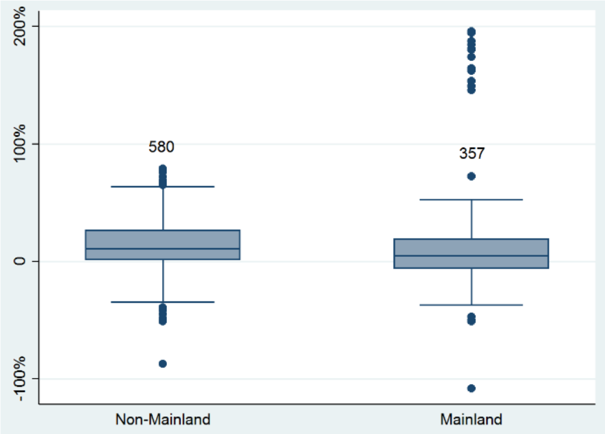
5a. Publication type



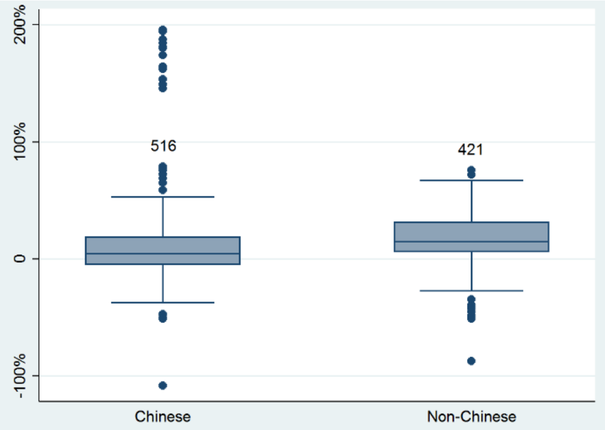
5b. Language of publication



5c. Mainland China institutional affiliations



5d. Chinese or non-Chinese authors



5e. Author affiliation types

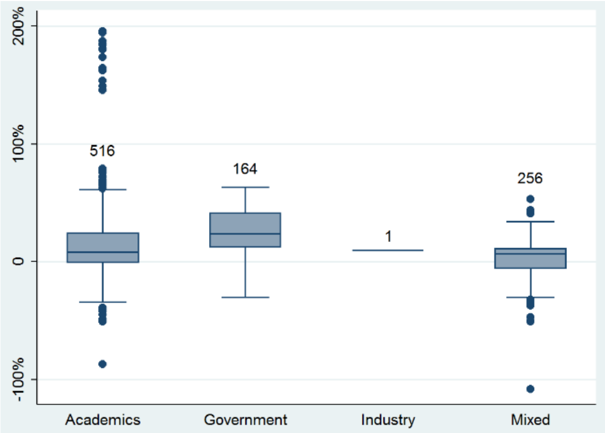


Figure 6a Box plots of RMB misalignment estimates: Subsample periods

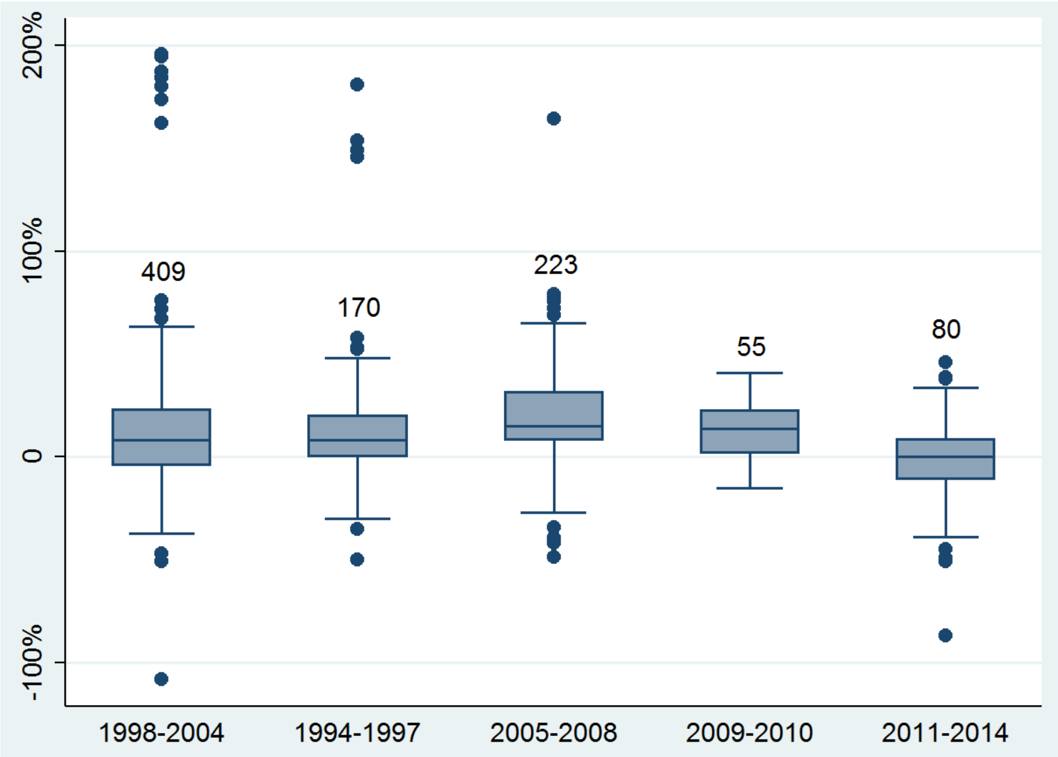


Figure 6b Dollar-based RMB exchange rate and averaged RMB misalignment estimate

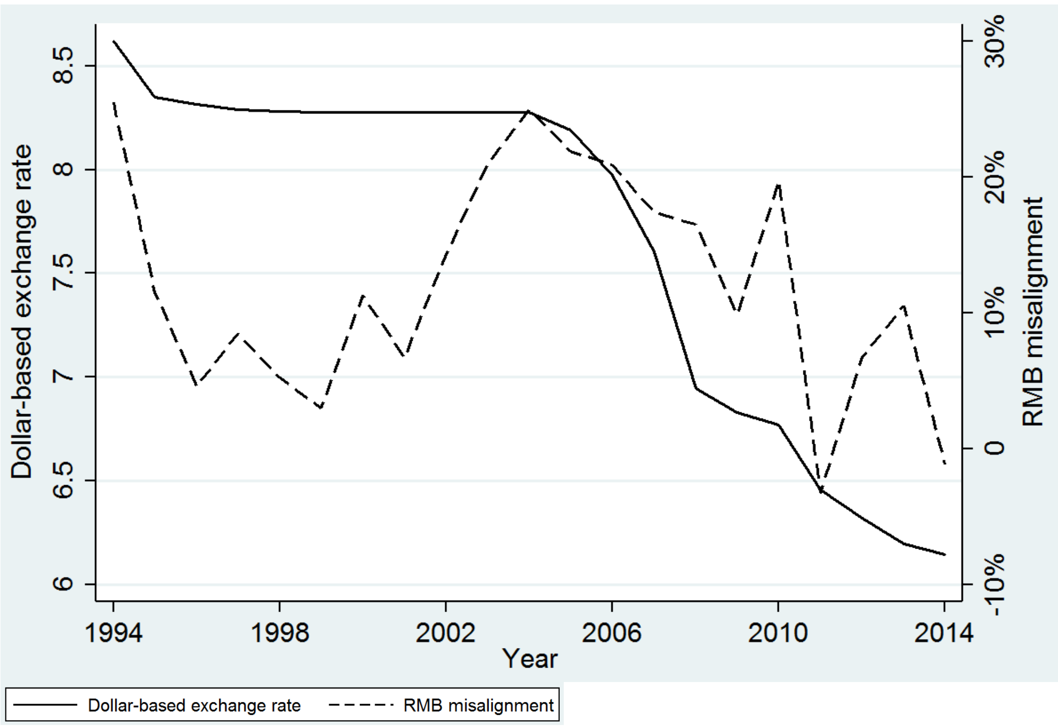
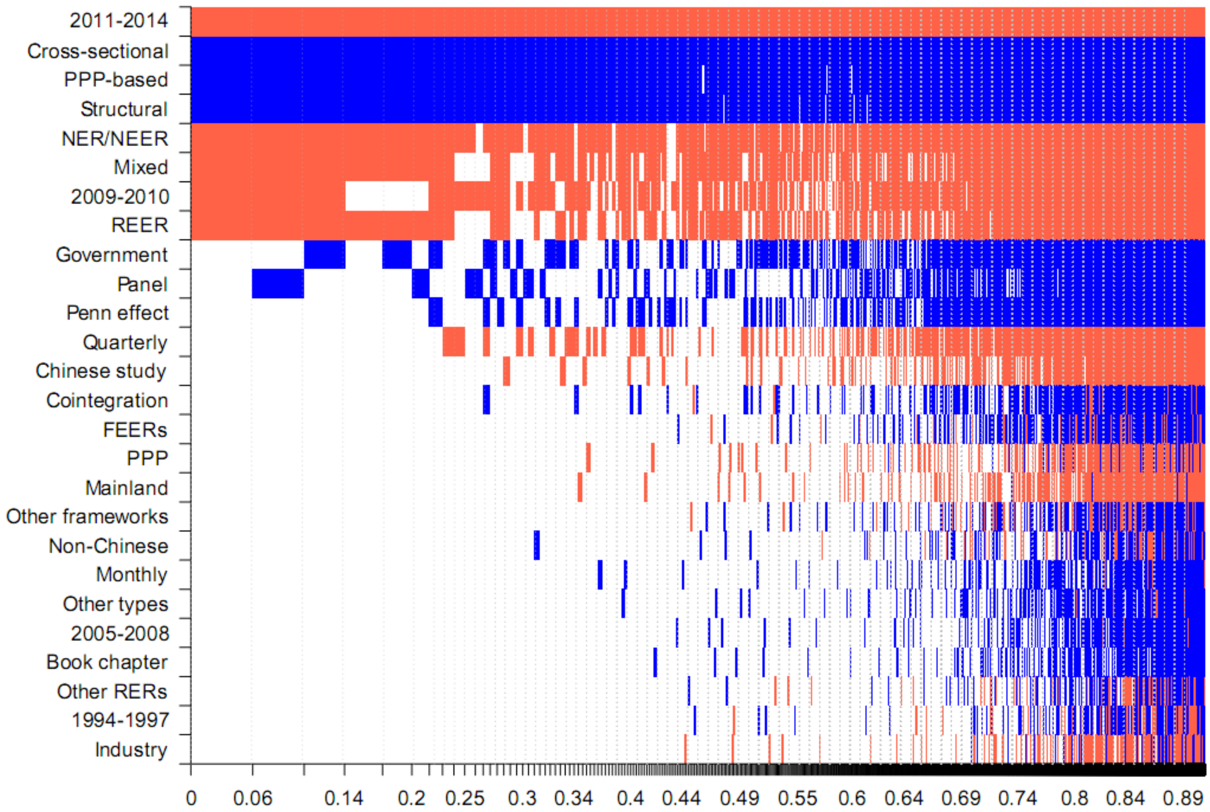


Figure 7 Top 6,000 models with the highest posterior model probabilities based on 26 study characteristic types



Note: The 26 study characteristic types (moderator variables) are listed on the vertical axis in descending orders of their PIPs. Each column represents a model specification with the column width indicating its posterior model probability, i.e. measure of the degree it is favored by the 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. 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 Density plot of the RMB misalignment estimate under the median probability model specification

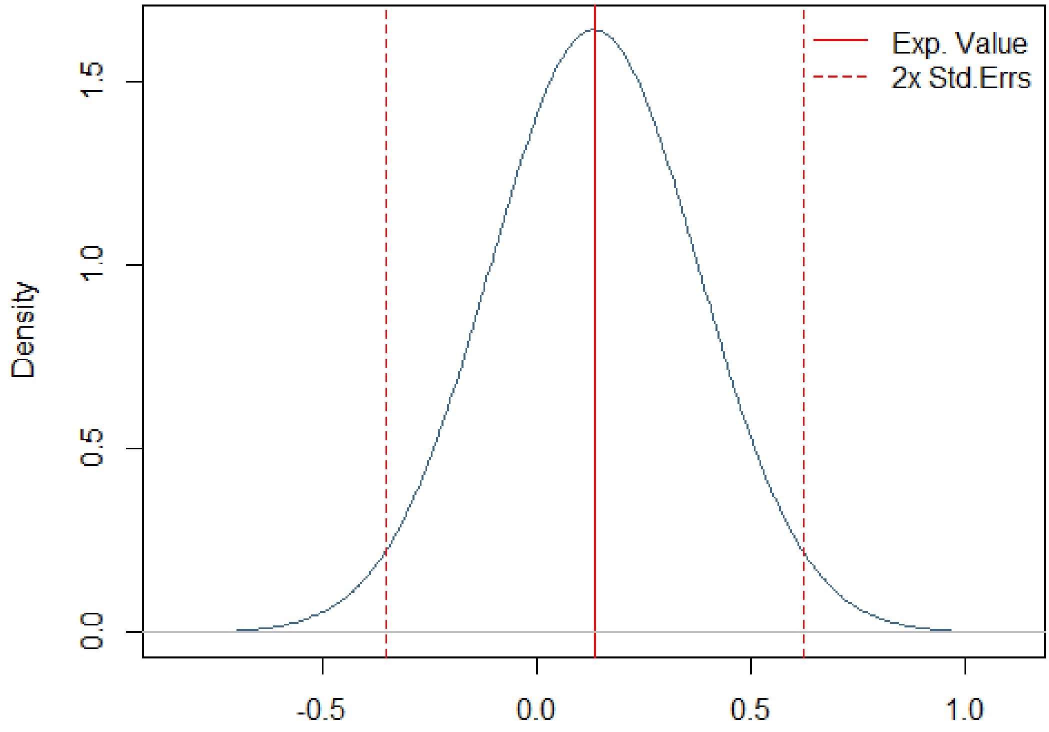
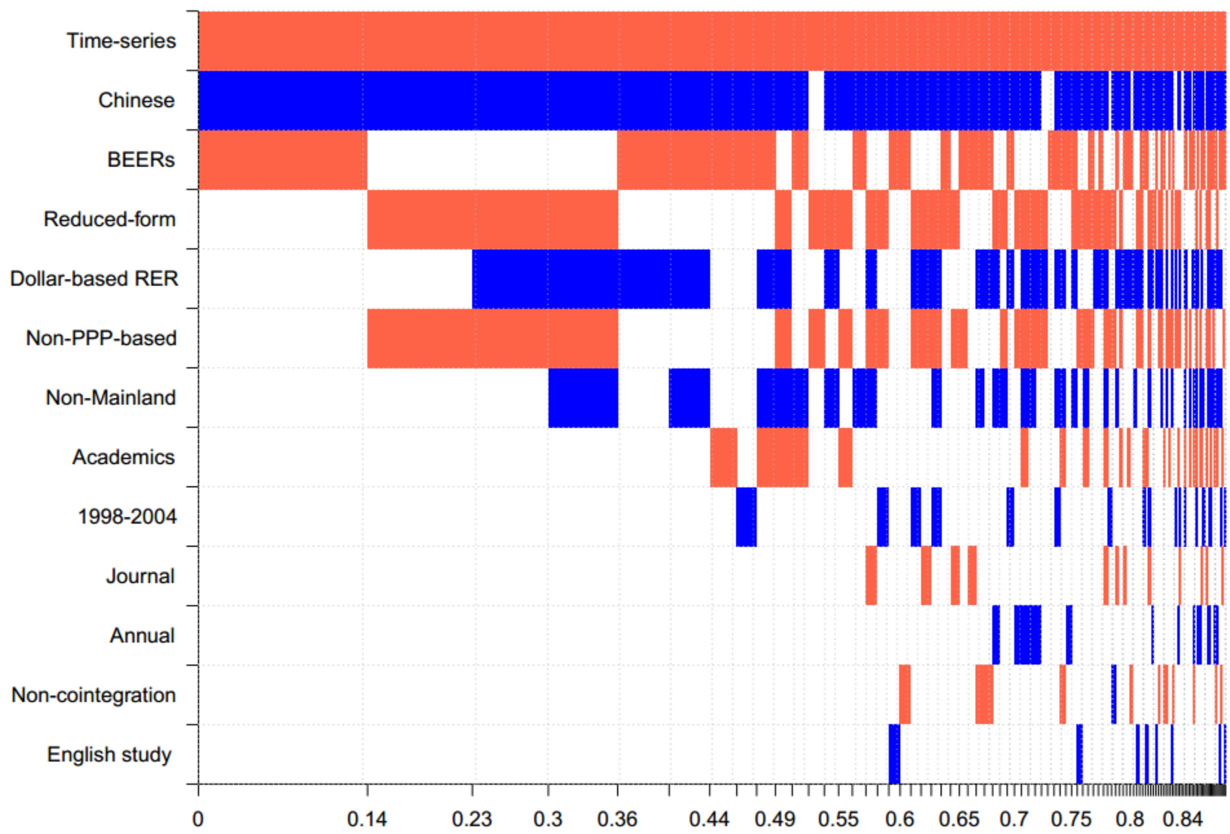


Figure 9 Top 100 models with highest posterior model probabilities (based on 13 study characteristic types)



Note: See the note in Figure 7. The 13 study characteristic types (moderator variables) are listed on the vertical axis in descending order of their PIPs. 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.

Appendices

A1 Sample of studies

ID	Study	Publication type	Language
1	Agya and Jun (2015)	Journal article	English
2	Aflouk, Jeong, Mazier and Saadaoui (2010)	Journal article	English
3	Almas, Grewal, Hvide and Ugurlu (2017)	Journal article	English
4	Benassy-Quere and Lahreche-Revil (2008)	Journal article	English
5	Benassy-Quere, Bereau and Mignon (2009)	Journal article	English
6	Benassy-Quere, Lahreche-Revil and Mignon (2011)	Journal article	English
7	Chang (2007)	Journal article	English
8	Chang (2008)	Journal article	English
9	Chang and Qin (2004)	Journal article	English
10	Chen (2009)	Journal article	English
11	Cheung, Chinn and Fujii (2007)	Journal article	English
12	Cheung, Chinn and Fujii (2009)	Journal article	English
13	Cheung, Chinn and Fujii (2010)	Journal article	English
14	Cheung, Chinn and Fujii (2017)	Journal article	English
15	Chou and Shih (1998)	Journal article	English
16	Christoph and Hossfeld (2014)	Journal article	English
17	Coudert and Couharde (2007)	Journal article	English
18	Cui (2013)	Journal article	English
19	Frankel (2006)	Journal article	English
20	Funke and Rahn (2005)	Journal article	English
21	Gan, Ward, Su and Cohen (2013)	Journal article	English
22	Garroway, Hacibedel, Reisen and Turkisch (2012)	Journal article	English
23	Giannellis and Koukouritakis (2018)	Journal article	English
24	Hall, Kenjegaliev, Swamy and Tavlas (2013)	Journal article	English
25	Hu and Chen (2010)	Journal article	English
26	Lipman (2011)	Journal article	English
27	Lü (2007)	Journal article	English
28	Nouira, Plane and Sekkat (2011)	Journal article	English
29	Peng, Lee and Gan (2008)	Journal article	English
30	Schroder (2013)	Journal article	English
31	Yang and Bajeux-Besnainou (2006)	Journal article	English
32	Yi (2010)	Journal article	English
33	You and Sarantis (2011)	Journal article	English
34	You and Sarantis (2012a)	Journal article	English
35	You and Sarantis (2012)	Journal article	English
36	Zhang and Chen (2014)	Journal article	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

ID	Study	Publication type	Language
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 article	English
52	Cline (2013b)	IIE article	English
53	Cline (2014a)	IIE article	English
54	Cline (2014b)	IIE article	English
55	Cline and Williamson (2008)	IIE article	English
56	Cline and Williamson (2009)	IIE article	English
57	Cline and Williamson (2010a)	IIE article	English
58	Cline and Williamson (2010b)	IIE article	English
59	Cline and Williamson (2011)	IIE article	English
60	Cline and Williamson (2012a)	IIE article	English
61	Cline and Williamson (2012b)	IIE article	English
62	Subramanian (2010)	IIE article	English
63	Shi and Yu (2005)	Journal article	Chinese
64	Sun and Sun (2013)	Journal article	Chinese
65	Wang (2015)	Journal article	Chinese
66	Wang and Cai (2007)	Journal article	Chinese
67	Wang and Lin (2013)	Journal article	Chinese
68	Wang and Yao (2008)	Journal article	Chinese
69	Zhang (2000)	Journal article	Chinese

Note: IIE refers to Peterson Institute for International Economics.

A2 Percentage misalignment estimates: Descriptive statistics

Year	Obs.	Mean	Std. Dev.	Min.	Max.
1994	49	25.50%	26.47%	-4.00%	181.00%
1995	39	11.55%	26.06%	-17.35%	153.60%
1996	39	4.58%	28.60%	-50.00%	145.80%
1997	43	8.51%	27.74%	-30.00%	149.30%
1998	43	5.16%	30.47%	-35.00%	162.40%
1999	54	3.03%	30.66%	-51.00%	173.90%
2000	58	11.28%	31.62%	-36.00%	180.20%
2001	61	6.69%	35.53%	-108.00%	187.50%
2002	64	14.30%	28.79%	-11.26%	194.60%
2003	59	21.01%	28.94%	-6.86%	196.00%
2004	70	25.01%	29.48%	-13.26%	184.40%
2005	87	21.88%	27.42%	-48.70%	164.60%
2006	47	20.94%	18.69%	-5.10%	75.80%
2007	37	17.45%	17.95%	-10.00%	79.00%
2008	52	16.53%	19.95%	-42.00%	78.10%
2009	42	9.94%	14.42%	-15.10%	40.20%
2010	13	19.65%	14.91%	-3.85%	40.70%
2011	51	-3.29%	24.14%	-87.00%	46.00%
2012	7	6.74%	15.14%	-11.40%	38.10%
2013	8	10.49%	10.34%	2.00%	31.00%
2014	14	-1.24%	21.52%	-27.20%	33.30%

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

A3 Study characteristic types

Study characteristic types	Description
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-based①	=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.
Theoretical and estimation specifications	
BEERs②	=1 if a model from the family of behavioral equilibrium exchange rate models or the productivity approach is used.
FEERs③	=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.
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 never educated in mainland China.

Study characteristic types	Description
Publication attributes cont.	
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), 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.
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:
① We do not distinguish the vintages of ICP survey data here. Note, however, that the ICP data revisions have a pronounced implication for estimating currency misalignment (Cheung and Fujii, 2014).

② "BEERs" refers to the family of behavioral equilibrium exchange rate models that includes the standard behavioral equilibrium exchange rate model (BEER), the permanent equilibrium exchange rate model (PEER), the equilibrium real exchange rate model (ERER), the Goldman Sachs dynamic equilibrium exchange rate (GSDEER), and the productivity approach (Cheung, Chinn and Fujii, 2010).

③ "FEERs" refers to fundamental equilibrium exchange rate model (FEER), IMF macroeconomic balance approach (MB), and the nature rate of exchange approach (NATREX). The models are theoretically quite similar.

A4 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 $\boldsymbol{X} = (X_1 . X_2 \ldots X_K)$ 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.

Which of the potential explanatory variables 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 $(X_1 . X_2 \ldots X_K)$ 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) alleviates the problem of model selection uncertainty by considering all possible model specifications and making 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 $p(M_k \mid y, \boldsymbol{X}_k)$ indicates how well a model fits the data. 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 PIP 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 (PIP) 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 PIP is a measure to assess the (relative) level that the data favor the inclusion of variable X_j in the chosen model. If the PIP 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 PIP 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 researcher knows 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). These are quite conservative and reflect unknown true model size and parameter signs.[26] Specifically, with 2^K possible models, a uniform model prior sets the common prior model probability to $p(M_k) = 2^{-K}$. The unit information prior on Zellner's g-prior sets the $g = N$ (N= 937, the number of observations in this exercise).

It is usually inefficient or unfeasible to compute all potential models as enumerating all models can be quite time intensive, especially with a 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

[26] There are alternative choices for priors on models and priors on parameters, including the beta-binomial model prior and benchmark prior on Zellner's g-prior (Ley and Steel, 2012). We tried these alternative choices and obtained results qualitatively similar to those reported in the main text.

Metropolis-Hasting algorithm that can go through the most important models with high posterior model probabilities (Zeugner and Feldkircher, 2015).

As an example of the Metropolis-Hastings algorithm,[27] the algorithm would first consider a model M_i and calculate 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 replaces M_i and the process continues. The distribution of posterior model probabilities converge when the number of repeated steps is sufficient. In the current study, all BMA computations use 1,000,000 burn-ins and 2,000,000 iterations to ensure a good degree of convergence.

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

BOFIT Discussion Papers

A series devoted to academic studies by BOFIT economists and guest researchers. The focus is on works relevant for economic policy and economic developments in transition / emerging economies.

BOFIT Discussion Papers
http://www.bofit.fi/en • email: bofit@bof.fi
ISSN 1456-4564 (print) // ISSN 1456-5889 (online)