行政院國家科學委員會專題研究計畫 期中進度報告

貨幣、匯率與動態均衡之學術前沿研究--子計畫七：匯率預測：估計風險之角色(2/4)
期中進度報告(完整版)

計 畫 主 持 人 ： 郭炳伸

報 告 附 件 ： 國外研究心得報告

處 理 方 式 ： 本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中 華 民 國 97 年 01 月 25 日

※大學學術追求卓越發展延續計畫執行報告格式

Explanation for the Form of the <u>Annual/Midterm/Final</u> Report "Program for Promoting Academic Excellence of Universities（Phase II）"

※　The <u>Annual/Midterm/Final</u> Report contains the following sections:

I		COVER
II	FORM1	BASIC INFORMATION OF THE PROGRAM
III	FORM2	LIST OF WORKS, EXPENDITURES, MANPOWER, AND MATCHING SUPPORTS FROM THE PARTICIPATING INSTITUTES（REALITY）.
IV	FORM3	STATISTICS ON RESEARCH OUTCOMES OF THIS PROGRAM
V	FORM4	EXECUTIVE SUMMARY ON RESEARCH OUTCOMES OF THIS PROGRAM
VI	APPENDIX I	MINUTES FROM PROGRAM DISCUSSION MEETINGS
VII	APPENDIX II	1.　PUBLICATION LIST（CONFERENCES, JOURNALS, BOOKS, BOOK CHAPTERS, etc.） 2.　PATENT LIST 3.　INVENTION LIST 4.　LIST OF WORKSHOPS/CONFERENCES HOSTED BY THE PROGRAM 5.　LIST OF PERSONAL ACHIEVEMENTS OF THE PIs 6.　LIST OF TECHNOLOGY TRANSFERS 7.　LIST OF TECHNOLOGY SERVICES
VIII	APPENDIX III	LIST OF PUBLICATIONS IN "TOP" JOURNALS AND CONFERENCES
IX	APPENDIX IV	SLIDES ON SCIENCE AND TECHNOLOGY BREAKTHROUGHS (TWO SLIDES FOR EACH BREAKTHROUGH)
X	APPENDIX V	SELF-ASSESSMENT

(Add extra lines or columns if needed.)

Program for Promoting Academic Excellence of Universities（Phase II）

Mid-Term   Report

匯率預測：估計風險之角色 (2/4)

Understanding Exchange Rate Predictability: The Role of Estimation Risk (2/4)

Serial number: NSC 95-2752-H-004 -002 -PAE

Overall Duration: Month ___04___ Year _2006_____ - Month _03_____ Year__2010____
Report Duration: Month __04___ Year __2007____ Month _03____ Year__2008____

National Chengchi University

Jan., 2008

II. (FORM1)   BASIC INFORMATION OF THE PROGRAM

<table>
<tr><td colspan="2">Program Title:</td><td colspan="5">匯率預測：估計風險之角色 (2/4)
Understanding Exchange Rate Predictability: The Role of Estimation Risk (2/4)</td></tr>
<tr><td colspan="2">Serial No.: NSC 95-2752-H-004 -002 –PAE</td><td>Affiliation</td><td colspan="3">(in English & Chinese)</td></tr>
<tr><td rowspan="4">Principal Investigator</td><td>Name</td><td>郭炳伸 Biing-Shen Kuo</td><td rowspan="4">Program Coordinator</td><td>Name</td><td colspan="2">(in English & Chinese)</td></tr>
<tr><td>Tel:</td><td>(02) 29393091 ext 81029</td><td>Tel:</td><td colspan="2"></td></tr>
<tr><td>Fax:</td><td>(02) 29387699</td><td>Fax:</td><td colspan="2"></td></tr>
<tr><td>E-mail</td><td>bsku@nccu.edu.tw</td><td>E-mail</td><td colspan="2"></td></tr>
</table>

	Expenditures[1] (in NT$1,000)		Manpower[2]:Full time/Part time(Person-Months)	
	Projected	Actual	Projected	Actual
FY 2006	1052	664	3	3
		-		-
		-		-
		-		-
Overall	1052	664	3	3

Serial No.	Project Title	Principal Investigator	Title	Affiliation
Sub-Project 7	匯率預測：估計風險之角色 Understanding Exchange Rate Predictability: The Role of Estimation Risk	郭炳伸 Biing-Shen Kuo	教授 Professor	國立政治大學 國際貿易學系 National Chengchi Univ.

Notes: [1,2] Please explain large differences between projected and actual figures.

Program Director/Principle Investigator Signature: Biing-Shen Kuo

III. (FORM 2) LIST OF WORKS, EXPENDITURES, MANPOWER, AND MATCHING SUPPORTS FROM THE PARTICIPATING INSTITUTES（REALITY）.

Serial No.: **NSC 95-2752-H-004 -002 -PAE**		Program Title: 匯率預測：估計風險之角色 **Understanding Exchange Rate Predictability: The Role of Estimation Risk**										
Research Item (Include sub projects)	Major tasks and objectives	Expenditures (in NT$1,000)					Manpower (person-month)					Matching Supports from the Participating Institutes (in English & Chinese)
		Salary	Seminar/ Conference-related expenses	Project- related expenses	Cost for Hardware & Software	Total	Principal Investigators	Consultants	Research/ Teaching Personnel	Supporting Staff	Total	
Sub-project 7	Develop of new estimators and explore its applications	516	148	0	0	664	1	0	3	0	4	0
SUM		516	148	0	0	664	1	0	3	0	4	0

IV. (FORM 3) STATISTICS ON RESEARCH OUTCOME OF THIS PROGRAM

LISTING		TOTAL	DOMESTIC	INTERNATIONAL	SIGNIFICANT[1]	CITATIONS[2]	TECHNOLOGY TRANSFER
PUBLISHED ARTICLES	JOURNALS						
	CONFERENCES						
	TECHNOLOGY REPORTS						
PATENTS	PENDING				-		
	GRANTED				-		
COPYRIGHTED INVENTIONS	ITEM						
WORKSHOPS/CONFERENCES[3]	ITEM	3	1	2			
	PARTICIPANTS	Around 60	Around 20	Around 40			
TRAINING COURSES (WORKSHOPS/CONFERENCES)	HOURS						
	PARTICIPANTS						
PERSONAL ACHIEVEMENTS	HONORS/ AWARDS[4]						
	KEYNOTES GIVEN BY PIS						
	EDITOR FOR JOURNALS						
TECHNOLOGY TRANSFERS	ITEM						
	LICENSING FEE						
	ROYALTY						
INDUSTRY STANDARDS[5]	ITEM						
TECHNOLOGICAL SERVICES[6]	ITEM				-	-	-
	SERVICE FEE				-	-	-

[1] Indicate the number of items that are significant. The criterion for "significant" is defined by the PIs of the program. For example, it may refer to Top journals (i.e., those with impact factors in the upper 15%) in the area of research, or conferences that are very selective in accepting submitted papers (i.e., at an acceptance rate no greater than 30%). Please specify the criteria in Appendix IV.

[2] Indicate the number of citations. The criterion for "citations" refers to citations by other research teams, i.e., exclude self-citations.

[3] Refers to the workshop and conferences hosted by the program.

[4] Includes Laureate of Nobel Prize, Member of Academia Sinica or equivalent, fellow of major international academic societies, etc.

[5] Refers to industry standards approved by national or international standardization parties that are proposed by PIs of the program.

[6] Refers to research outcomes used to provide technological services, including research and educational programs, to other ministries of the government or professional societies.

V. (FORM4)   EXECUTIVE SUMMARY ON RESEARCH OUTCOMES OF THIS PROGRAM

(Please state the followings concisely and clearly)

1.   General Description of the Program: Including Objectives of the Program

The research attempts to offer econometric explanations to the near random-walk exchange rates. It argues that previous empirical evidence for or against predictability in exchange rate movements might have been considerably flawed by the existence of estimation risk due to the strong persistence in fundamentals. The primary goal of the project is in a pursuit of a more reliable inference procedure for the predictability both in-sample and out-of-sample by appropriately controlling the estimation risk.

To achieve the goal, an averaging estimator that combines information optimally both from the univariate time series under study and from cross-sectional time series is developed.

2.   Breakthroughs and Major Achievements

Evidence for the exchange rate predictability in the past literature has been mixed. In contrast, the current project, after controlling the estimation risks, is able to establish a more uniform evidence for the predictability, whether the forecasting horizons are short or long. This is somehow remarkable because to our knowledge, little evidence for the exchange rate predictability in the short horizons was found in the literature.

In addition to establishing evidence for the predictability, two major conclusions emerging from the research so far can be summarized: 1) the magnitude of the estimation risks is so high that the exchange rate predictability can be masked even when it exists in the data; 2) The information about exchange rate movements from cross-section is valuable in the ways that it can reduce the estimation risk and thus improve the testing power for predictability, if it could be exploited effectively as the averaging estimator does.

We are hoping in the future to analyze the statistical property of the averaging estimator better by establishing the associated asymptotic theory. This will consist of another major achieve achievements for the project.

3.   Categorized Summary of Research Outcomes. The criteria for top conferences and journals should be given and

introduced briefly in the beginning of this section. In each research area, please give a brief summary on the

research outcomes associated with the area. Note that the summaries should be consistent with the statistics

given in Form3. Please list and number each research outcomes in sorted order in Appendix II, and list all the

publications in top conferences and journals in Appendix III.

3.1 Development of an avaraging estimator:

An averaging estimator that is to control potential estimation risks associated with the predictive regression is developed in the first-year study. The sources of the estimation risks comes from high persistence of predictive regressors, and the dependent variable being the overlapping sums of short-horizon change in log exchange rate. The former creates bias in small-samples and the latter brings forth remarkable estimation variability in long-horizon predictions.

The considered averaging estimator optimally combines two alternative estimators that differ in their bias and precision characteristics. By construction, the suggested estimator for the slope coefficients utilizes information from cross-sections in a similar way that the panel-based estimators do. The implicit assumption underlying the use of information from cross-sections for our estimator, however, is very much different from that for the

panel-based estimators. In contrast, the panel-based estimators are built on the assumption that the slope coefficients are all the same for all the cross-sectional countries. On the other hand, the averaging estimator allows for separate slope estimate for each cross-section country as the OLS estimator does, but makes use of the cross-sectional information that the OLS estimator does not. Thus both the averaging estimator and the pooled estimator are the same to reduce the estimation errors, but differ in the way how the cross-sectional information is processed. Yet, our averaging estimator has the advantages of producing more reasonable slope estimates.

3.2 Risk reduction: simulation analysis

We examined to what extent the proposed estimator can improve over the traditional estimator in terms of risk reduction through simulations. Under the setup that mimics the reality, we documented that the averaging estimator empirically dominates the least-square (LS) estimator, regardless of which simulation scenario is considered. Virtually the risk reductions using the averaging estimator can be as large as between 10\% and 35\%, compared with the LS estimator. More importantly, the risk improvement by the averaging estimator is embodied further into power gains in testing. Our simulation shows that the power gains from using the averaging estimator, again relative to the LS estimator, is 10% to 30% or more in many cases. An significant implication of the finding is simply that the predictability alternative can now be better detected from the data when the test statistics are based on the averaging estimator.

3.3 A re-examination of the exchange rate predictability

We re-investigated the empirical validity of the exchange rate predictability applying the averaging estimator. The testing strategy basically follows that utilized in the literature where these studies all base their inference on the bootstrap approach in order to control for small-sample bias for which the asymptotic approximation generally fails to correct.

We accessed the relative forecast accuracy of the two competing models with Theil's U and DM statistics. It should be noted that the problem with estimating the long-run variance precisely when calculating the DM statistic often leads to spurious inference. Important messages emerging from the empirical exercises include:

1) There is now much more significant evidence presented for the dominance of the monetary model over the random walk when predicting, after accounting for estimation risks using the considered estimator. With only few exceptions, the p-values associated with the averaging estimator for both statistics are smaller, relative to those associated with the LS estimator.

2) It stands out from the results that controlling over the risks uncovers more favorable evidence in supports of the monetary model, while there is essentially no evidence for so when leaving the risks unattended. Many more instances of this are found from the reported Theil's U statistic. Particularly, at almost all horizons, the monetary model is found to be superior to the random walk in terms of predictability for Germany and Japan. This contrasts sharply with the previous findings where little evidence for predictability is reported. Considering the Theil's U statistic is more robust, this evidence lends quite a good deal of credence to the predictability at both short- and long-horizons.

3.4 Asymptotic theory of the averaging estimator

The use of the averaging estimator in testing for exchange rate predictability brings forth some econometric interesting questions. This entails the development of an asymptotic theory of the averaging estimator. We invoke a local-to-unity framework to build the asymptotic theory based on the observation with inherent high persistence in the data. We are now able to derive the asymptotic distributions of the averaging estimator under the simplified assumptions where regression errors are uncorrelated with predicting variables. The asymptotic distribution derived is a mixture normal. The mixture normal collapses into the limit distribution of the least-square estimator, or that of the panel estimator, when either receives zero weights in forming the averaging estimator.

4. Program Management: the Mechanism for Promoting Collaboration and Integration among the Institutes Involved

The mini conference held at Academia Sinica in Jan 2007 exposed me to the ideas contributed by other principle investigators of different sub-projects. Admittedly I learned quite a bit from these team partners. Some of the work and results were intriguing. Importantly, I found that there should have more links to each other among sub-projects than it used to project.

We shall have another mini conference in similar form to be held in the coming February, 2008. I am looking forward to the chance to share my research efforts with other investigators and constructive comments from them.

5. A Summary of the Post-Program Plan（Including the Detailed Description of Budget and Plan Adjustment of the

    next year）

My ongoing research efforts will give to the completion of the asymptotic theory of the averaging estimator. The current asymptotic results are obtained under some simplified assumptions that are hard to meet in realistic applications where regression errors are usually autocorrelated and correlated with predictors. The asymptotic theory of the averaging estimator however would be even more difficult to derive in this context, because of endogeneity resulting from the correlations. To obtain the asymptotic mixture normal for the averaging estimator under more realistic assumptions would be involved with some bias-correction mechanism as the previous literature on the unit root theory. An inference theory for the averaging estimator will be to build, following the development of the asymptotic theory. These would be the main takes for the project of the year and the one to come.

6. International Cooperation Activities (Optional)

Prof. Bruce Hansen of University of Wisconsin at Madison has been working on the averaging estimator around the same time as the research got started. One of his first papers along the line got published in a recent issue of Econometrica, marking research on the averaging estimator a potential important direction to move in the future. Possible research collaborations with him in terms of visiting him would greatly help improve the exposition of my research.

VI. APPENDIX I: MINUTES FROM PROGRAM DISCUSSION MEETINGS

VII. APPENDIX II:

1. **PUBLICATION LIST（CONFERENCES, JOURNALS, BOOKS, BOOK CHAPTERS, etc.）**

    DOING JUSTICE TO FUNDAMENTALS IN EXCHANGE RATE FORECASTING: A CONTROL OVER ESTIMATION RISKS, PRESENTED IN NORTH AMERICAN SUMMER MEETING OF ECONOMETRIC SOCIETY, JUNE 2007, DUKE UNIVERSITY, USA; AND EUROPEAN SUMMER MEETING OF ECONOMETRIC SOCIETY, AUGUST 2007, BUDAPEST, HUNGARY.

2. **PATENT LIST**
3. **INVENTION LIST**
4. **LIST OF WORKSHOPS/CONFERENCES HOSTED BY THE PROGRAM**
5. **LIST OF PERSONAL ACHIEVEMENTS OF THE PIS**
6. **LIST OF TECHNOLOGY TRANSFERS**
7. **LIST OF TECHNOLOGY SERVICES**

VIII. APPENDIX III: LIST OF PUBLICATIONS IN "TOP" JOURNALS AND CONFERENCES

IX. APPENDIX IV: SLIDES ON SCIENCE AND TECHNOLOGY BREAKTHROUGHS

(TWO SLIDES FOR EACH BREAKTHROUGH)

Program for Promoting Academic Excellence of Universities（Phase II）
<u>Midterm</u> ASSESSMENT

PROGRAM/**Sub-project** TITLE: DOING JUSTICE TO FUNDAMENTALS IN EXCHANGE RATE FORECASTING: A CONTROL

OVER ESTIMATION RISK

	ASSESSMENT SUBJECT	SCORE (1~5, LOW TO HIGH)
PROGRAM'S CONTENTS & PERFORMANCE	Importance & Innovation of the Program's Major Tasks	5
	Clarity and Presentation of the Report	4
	Viability of the Program's Approaches & Methodologies	5
	Principal Investigator's Competence for Leading the Program	5
	Interface & Integration between Overall & Sub-Project(s)	4
	Interface & Integration among All Sub-Projects	5
	Manpower & Expenditures	5
PROGRAM'S RESULTS	Contribution in Enhancing the Institute's International Academic Standing	5
	Impact on Advancing Teaching or on Technology Development	5
Total Score		43

REVIEWER'S COMMENTS & SUGGESTION:

This project addresses an important question concerning whether fundamentals can predict exchange rate movements. The author proposes a Stein-like shrinkage estimator to tackle the exchange rate predictive regression problem. The simulation results point that, as compared to the usual OLS estimator, the proposed shrinkage estimator performs well in terms of risk reduction. The application of this shrinkage estimator also indicates that the monetary fundamentals can beat the simple random walk model in forecating exchange rate at short horizons out-of-sample. The followings are my comments about this project.

1. The author might need to explain the relationship between bias, precision, and risk in the beginning of the paper, because the readers of this paper should be empirical-oriented. Furthermore, the author should explain the wording `estimation error' and `bias estimation'. For example, p.1 (line -1), I am confused with the sentence `--- the bootstrap helps reduce the bias, it introduces additional errors due to the bias estimation.'

2. P.4. x(it) is assumed to be mean-reverting by the author. So it is I(d) with d<1. However, in (7) of page 4, (1-L)x(it) is used as dependent variable, indicating that this dependent variable is not a short memory process. This observation is important because the VEC representation displayed in (6) and (7) is problematic under this circumstance.

3. To make VEC representation balanced in (6) and (7), there should be more terms in the right-hand side of (6). The author needs some works to justify this kind of arrangement.

4. If x(it) are mean-reversting, and e(it) are I(1), then the econometric modeling behind these two variables should be more elaborate. This task might also serve as another possible explanation for the puzzling behavior of the exchange rates.

5. It is very interesting if the author can estimate the covariance matrix between the OLS and the grand average.

6. P.5, I suggest the author follow the setting in (1) of Berkowitz and Giorgianni (2001) in writing up his equation (8) to make the long-horizon predictive regression framework more evident.

7. I also hope the author can help me understand why his paper can offer econometric explanations to the near random walk exchange reates as argued in page 1. I thought the author take the I(1) behavior of fundamentals as given. Of course, it is highly likely that I misunderstand the intention of the author outlined in the introuction section.

8. The author claims the grand average estimator has larger bias. Why?

9. The references do not contain the paper of Phillips (1998).

10. Can you explain more clearly about the the source and procedure of Campbell-Perron.

11. Can the author use % to present the power performance of the shrinkage estimator. The author may need to write down the models for comparison more clearly in Section 5.2.

The reviewer made a few comments and suggestions that are particularly helpful in revising and improving the paper. Below is my short reponses to his comments.

1. **In the paper, the estimation risks include 2 parts, bias and estimation uncertainty. The latter is interchanged with estimation errors. Either estimation error or uncertainty refers to as the estimation efficiency. The sentence that the reviewer is concerned with, 'the bootstrap helps reduce the bias, it introduces additional errors due to the bias estimation' should be re-written. The current writing indeed lead to a big confusion.**

2. **The independent variable or the depedent variable in the predictive regressions that the paper is looking at is not of long memory process. They do have the high persistence nature which is similar to long memory process. But their asymptotics are quite different. In particular, the variables of concern are best described as the processes that are near unit roots. Under these specification, equations (6) and (7) or the VECM is subject to no errors.**

3. **The expressions in equations (6) and (7) indeed are not make clear. I shall add to the paper in the revision a footnote to elaborate how the VECM can be derived.**

4. **Our empirical observations show that the residual terms from the predicitve regression are I(0) in general, implying that error term, e(it), might follow an I(0) process as well.**

5. **This is a great suggestion. I shall calcualte the correlation between grand mean average and the old estimates, and report it in the revision. The quantity should offer some important information about the quality of the combined estimators.**

6. **Berkowitz and Giorgianni (2001) gave a good account of the derviations of the long predictive regression. To save the space, the current version of the paper did not give it a detailed account of equation (8). Again, a footnote will be provided to make sense the equation.**

7. **The motivation in the paper was a bit mislieading. It was to offer an econometric explanation to why exchange rate have been hard to be predictable, given that they are near unit root nature. Therefore, it is not to explain why they are near unit root, but explain their little predictability. The introduction will be adjusted and rewritten to make clear the motivation to aovid confusions.**

8. **The grand average has larger bias because it is compared to any single coefficient estimate. Under the alternative where predictability exists and each coefficient is not equal to zero. The sum of the coefficent estimates will be likely to be larger than any single coefficient estimate, if the bias is interpreted as a distance from zero, the null of the research. I shall re-work on this and have it revised.**

9. **Phillips (1998) will added to the reference accordingly.**

10. **Campbell and Perron procedure is known to applied research. To save space I did not explain its details, but the sources will be listed in the reivsion.**

11. **The benchmark for the comparison given in Sec. 5.2 is the models estimated by OLS. And the comparison there is based on the realizations of the underlying data generating processes. So it will be impossible to offer any power comparsions because we would never come to know the underlying DGP, thereby can not drawing samples repeatedly.**

Program Reviewer's Signature: <u>蔡文禎　Tsai, Wen-Jen</u>

Notes: The program reviewers are invited by the National Science Council.

Doing Justice to Fundamentals in Exchange Rate Forecasting: A Control over Estimation Risks

Jan., 2007

Abstract

A shrinkage estimator that accounts for estimation risks is developed and employed to re-access whether the monetary fundamentals help predict changes in exchange rates. While the estimation risk in term of mean squared errors takes place in the presence of parameter uncertainties, it is more problematic in the context of the exchange rate predictive regression where the time series under study are typically short and the predictors are highly persistent. Monte-Carlo simulations clearly demonstrate that comparing to the least-square estimator, the magnitude of estimation errors associated with our shrinkage estimator is 10% to 35% less. Moreover, the risk reductions convert into sizable power gains for tests for predictability, yielding a robust inference. In contrast to the previous studies, a uniform evidence of the higher ability of monetary fundamentals to forecast exchange rate movements is now found at short horizons out-of-sample.

Keywords: random walk, fundamentals, exchange rate forecasting, shrinkage estimator, estimation risk, bootstrap
JEL classification: C22, C52, C53, F31, F47

1 Introduction

Whether economic fundamentals help predict exchange rate movements remains an unsettled issue. Messe and Rogoff in a series of papers in 80s first demonstrated that a simple random walk forecast of the exchange rate outperforms alternative competing models based on economic theories. This is to say that economic fundamentals such as money supplies, interest rates and outputs simply contain little information in explaining future changes in exchange rates, an evidence for the failure of the economic models. Some subsequent studies do find success for fundamentals-related models that improve upon random walks out-of-sample. Yet, the results do not lend enough credence to shaking the puzzling fact that the exchange rates can be well-approximated as random walks, where their robustness remains to be established. While it is still difficult to attribute fundamentals to changes in exchange rates empirically, a recent theory of Engel and West (2004) justifies the near random-walk behavior of the exchange rate, even in the presence of a link between exchange rates and fundamentals.

Our study is motivated by one of the key assumptions on which the theory of Engel and West is built: Fundamental variables have autoregressive unit roots. We take the assumption as a natural beginning point of the study, and want to explore its empirical implications in the predictive regressions used in the literature. In contrast to economic theory of Engel and West, the paper attempts to offer econometric explanations to the near random-walk exchange rates, and argues that previous empirical evidence for or against predictability in exchange rate movements might have been considerably flawed by an existence of estimation risk due to the strong persistence in fundamentals. The primary goal of the paper is in a pursuit of a more reliable inference procedure for the predictability both in-sample and out-of-sample by appropriately controlling for the estimation risk.

The persistence nature in fundamentals has created subtle problems for estimation of and inference for the economic models. As one outstanding aspect of the problems, when fundamentals are highly persistent, the OLS estimates are very likely to subject to small-sample bias (Stambaugh, 1999). To correct for the bias, previous work appeals to the bootstrap techniques (for example, Mark, 1995; Kilian, 1999). Although the bootstrap helps reduce the bias, it introduces additional errors due to the bias estimation. On the

other hand, the persistence may lead to a substantial amount of variations in parameter estimation in the region of near unit root, implying an efficiency loss (Phillips, 1998). As a consequence of the loss, inferences for forecasting performance might well be biased in favor of random walks, even when the data are generated by the fundamentals-based models (Rossi, 2005). Nevertheless, it is well-known that there is a clear trade-off between bias and estimation errors in small samples. Thus, much previous work just provides partial answers to the predictability at issue by only looking at either bias or estimation errors. Little existing work however investigates how the tradeoff would impact on model estimations and forecasting inference.

The paper deals with both problems with small-sample bias and with estimation errors by proposing a shrinkage estimator. The shrinkage estimator is derived by minimizing mean-squared forecasting errors in the spirit of Judge and Mittelhammer (2004), and thus can properly take into account the tradeoff. Specifically, the proposed estimator is an optimal-weighted average of the OLS estimator from a certain exchange rate time series and the grand average of the OLS estimators from the panel of the exchange rates series under study. In this way, we exploit cross-sectional information as those using panel methods (Mark and Sul, 2001), but in quite a different manner. By shrinking the asymptotically unbiased OLS estimator full of estimation errors toward the biased grand average, the estimation risk can now be considerably reduced. Of empirical significance is that the use of the shrinkage estimator can potentially convert the estimation accuracy improvement into power gains for the forecasting ability testing. As a result, when exchange rate variations is predictable, the tests are ought to be more capable of detecting it, and the evidence by tests based on the mounted estimator would tend to be more trustworthy than that presented before. Our simulations clearly demonstrate the merits of the proposed semi-parametric estimator in reducing the estimation risk. Calibrating parameters to the data set that Kilian (1999) examined, our Monte-Carlo study shows that the power gains are found to be between 10% and 30% greater than those using OLS estimates. A technical innovation of our approach lies in an application of the bootstrap to parameter estimations, instead of plugging in consistent estimates as suggested by Judge and Mittelhammer (2004).

Applying the newly developed estimator, we re-access the central issue of the study: whether fundamentals can predict exchange rate movements. We also quantify potential

estimation risk which the previous research using conventional estimators could have been subjected to. Significantly, we find that the fundamentals-based models now consistently perform better than random walks, across various horizons. We conclude that the inability of the economic models to yield better forecasts might well be due to the estimation risk that has not been accounted for.

The remainder of the paper is organized as follows. Section 2 describes the monetary models of exchange rate determination and the corresponding predictive regression. In Section 3 we develop the shrinkage estimator used in the paper. Section 4 numerically analyzes the statistical advantages of the shrinkage estimator over the conventional estimator. Section 5 re-examines the exchange rate predictability, and presents the empirical findings. Finally, Section 6 concludes. Technical descriptions of the shrinkage estimator is left to the appendix.

2 A monetary model of exchange rates and its VAR representation

We consider a standard long-run models of exchange rates where purchasing power parity and uncovered interest parity are assumed to hold. Let e_{it} be the log nominal exchange rate at time t between currency i and the benchmark currency, denoted by '0'. The exchange rate is currency i price of a unit of currency 0. Let r_{it} be the nominal interest rate and p_{it} be the price level associated with currency i respectively. Thus,

$$r_{it} - r_{0t} = \mathrm{E}_t e_{it+1} - e_{it}, \tag{1}$$

$$e_{it} = p_{it} - p_{0t}, \tag{2}$$

where E_t is the conditional expectation operator.

When money markets, both domestic and foreign, are in equilibrium,

$$m_{it} - p_{it} = \lambda \, y_{it} - \phi \, r_{it}, \tag{3}$$

$$m_{0t} - p_{0t} = \lambda \, y_{0t} - \phi \, r_{0t}, \tag{4}$$

where m_{it}, y_{it} and p_{it}, respectively, are the log of money supply, output and price level of currency i. The analogous variables for currency 0 are m_{0t}, y_{0t}, and p_{0t}. Both money equations share common income elasticity λ and interest semi-elasticity ϕ. Combining (1),

3

(2), (3) and (4),

$$\mathrm{E}_t e_{it+1} - e_{it} = \mathrm{E}_t \triangle e_{it+1} = -\frac{1}{\phi}(f_{it} - e_{it}) = -\frac{1}{\phi}x_{it}, \tag{5}$$

where $f_{it} = (m_{it} - m_{0t}) - \lambda(y_{it} - y_{0t})$ is the monetary fundamental values, and thus $x_{it} = f_{it} - e_{it}$ is the deviation of the exchange rate from its fundamental value. The equivalence relation simply states that the exchange rate is expected to react to deviations from its fundamentals. The no bubble solution to the expectation difference equation (5) would deliver the present value model:

$$e_{it} = \frac{1}{1+\phi}\sum_{j=0}^{\infty}(\frac{\phi}{1+\phi})^j \mathrm{E}_t f_{it+j}.$$

Deducting f_{it} from both sides gives

$$e_{it} - f_{it} = \sum_{j=1}^{\infty}(\frac{\phi}{1+\phi})^j \mathrm{E}_t \triangle f_{it+j}.$$

These equations make it easy to explain a long-run equilibrium relation between exchange rates and the monetary fundamentals. As long as f_{it} is non-stationary processes, e_{it} is an I(1) process and $f_{it} - e_{it}$ is an I(0) process by the aforementioned relation. As a result, there is a cointegration between e_{it} and f_{it} with the pre-specified cointegrating vector (1,-1). Using the VAR framework, the cointegration implies a bivariate vector error correction representation for $y_{it} = (e_{it}, f_{it})'$: $y_{it} = c_0 + c_1 y_{it-1} + \cdots + c_p y_{it-p} + \varepsilon_{it}$. Some re-arranging can yield another vector error correction (VEC) model for e_{it} and x_{it}.[1] In particular, our econometric analysis centers on a VEC model that is frequently used in the exchange rate forecasting literature:

$$\triangle e_{it} = \beta_i^{(1)} x_{it-1} + \varepsilon_{it}^1, \tag{6}$$

$$\triangle x_{it} = \alpha_i x_{it-1} + \sum_{l=1}^{p_i} c_{il}^2 \triangle e_{it-1} + \sum_{l=1}^{p_i} d_{il}^2 \triangle x_{it-1} + \varepsilon_{it}^2, \tag{7}$$

where $(\varepsilon_{it}^1, \varepsilon_{it}^2)$ are assumed to be i.i.d. with a vector mean zero and a nonsingular covariance matrix. Equation (6) is the short-horizon predictive regression employed by Mark (1995), Berkowitz and Giorgianni (1997), Groen (1997), Kilian (1997), Berben and Dijk (1998) and Mark and Sul (2001) to study the exchange rate predictability. One of the goals of these studies as well as ours is to investigate if the mean reversion of x_{it} can improve the forecast

[1]This follows by noting that $\triangle x_{it} = \triangle f_{it} - \triangle e_{it}$.

4

accuracy by testing for $\beta_i^{(1)} = 0$ against $\beta_i^{(1)} > 0$. The positive sign of the slope coefficient $\beta_i^{(}1)$ suggests that e_{it} is expected to decline when above its fundamentals. Both equation (6) and (7) serve as data generating process in our bootstrapping algorithm after fitting them to the data. Because the log exchange rate returns to its fundamental value over time, more than 1-period ahead change in e_{it} may be characterized by deviations from the monetary fundamentals as well. Thus, the predictive regression (6) can in general be of the form:

$$e_{it} - e_{it-k} = \beta_i^{(k)} x_{it-k} + \varepsilon_{it}, \tag{8}$$

where k represents the forecast horizons and is set to 1, 4, 8, 12 and 16 in the analysis. Not to complicate the use of notations, the superscript (k) will be saved from the exposition following, unless there are confusions created.

Our analysis takes up two major related empirical concerns. As stated above, one would be to search for the empirical validity for the monetary model of exchange rates using our proposed estimator. By controlling for the estimation risk, the estimator is hoped to achieve a high degree of statistical accuracy in establishing the proposition. Moreover, the previous discussions suggest that for exchange rate to be predictable, exchange rates are ought to be cointegrated with the monetary fundamentals first. Therefore, another task is to investigate the existence of the cointegration between e_{it} and f_{it} based on the suggested estimator.

3 Econometric Framework

This section develops a shrinkage estimator that is to control potential estimation risks associated with the predictive regression. The motivation for a need for the estimator will be discussed immediately.

3.1 Sources of estimation risk

The OLS estimation of the slope coefficient tends to suffer more estimation risk than ordinary regressions. It results from some distinctive characteristics for the predictive regressions. First, the predictor x_{it}, or the deviation of exchange rate from fundamentals, is documented to be a highly persistently process. The high degree of persistence in forecasting variables, however, may generate a biased estimate of β_i in small samples, although do not alter the unbiasedness asymptotically. The magnitude of the bias is positively proportional to the

degree of persistence in forecasting variables for a given sample size, as demonstrated by Stambaugh (1999).

Second, the predictor possesses a much smaller variability than the dependent variable does (Mark and Sul, 2004). It is regarded that the small variance might not affect some statistical properties of the estimators asymptotically. But in finite samples, a small variance in regressors suggests a small signal-to-noise ratio, yielding imprecise estimates of β_i. It turns out that the imprecision could not be safely ignored when taking into account another key feature of predictive regressions as below.

Third, the dependent variable is long-horizon change in log exchange rate that amounts to overlapping sums of short-horizon change in log exchange rates. Valkanov (2003) have shown that in some parametric settings, the OLS estimator of β_i might be inconsistent, when the link of the small variance in predictors to the overlaps in dependent variables is considered. In other words, the small-sample distributions of the estimates of β_i is more dispersed with the forecast horizon k. This is simply another reflection of the imprecise estimation of β_i due to the small variance.

The estimation risk appears to take a form of either bias or inefficiency, whether it is due to the high persistence or the small variability. The little consensus emerging from empirical findings in the literature may come from the fact that econometric methods adopted can only deal with some parts of the risks. It is thus important to have a comprehensive approach that is capable of controlling for the overall estimation risks. This will be the subject to be studied in the next subsection.

3.2 A shrinkage estimator

To reduce the estimation risks in the context of predictive regressions, we consider a shrinkage estimator for β_i defined by

$$\tilde{\beta}_i = \omega_i \cdot \hat{\beta}_i + (1 - \omega_i) \cdot \bar{\beta},$$

where ω_i is the 'shrinking factor' that assigns weights to the two estimators in combination, $\hat{\beta}_i$ is the OLS estimator, and the grand average estimator is given by

$$\bar{\beta} = \frac{\sum_{i=1}^{N} \hat{\beta}_i}{N}$$

which takes the average of all the OLS estimators of the slope coefficients of the N countries. As it appears now, the shrinkage estimator considered is a Stein-like estimator which linearly combines two alternative estimators that differ in their bias and precision characteristics. Stein (1956) showed that shrinking sample means toward a fixed constant, under some conditions, can reduce estimation risks. The key is to create a tradeoff between bias and accuracy by introducing a biased estimator, represented by the fixed target, to combine with the unbiased estimator, represented by the sample mean. The statistical decision theory suggests the existence of an interior optimum in the tradeoff between bias and precision. Taking a proper weighted average of the unbiased and the biased estimators would constitute one of feasible approaches to attaining the optimal tradeoff. The idea of shrinkage has been widely applied to the problem of portfolio selection (for example, Jorion, 1985; Dumas and Jacquillat, 1990) and the context of prediction (Copas, 1983).

Our shrinkage estimator inherits the idea of Stein's (1956) seminal work. To see this, it is important to recognize that 1) the individual OLS estimator of the slope coefficient is unbiased asymptotically, but has lots of estimation errors due to the small variance of and high degree of persistence of the predictors, and that 2) the opposite is true for the grand average of all the cross-sectional slope estimators which is subject to larger bias but less estimation errors due to the averaging. Thus the intuition of the theory simply suggests that taking a weighted average of the OLS estimator and the grand mean estimator is likely to yield an estimator with less risk in the context of predictive regressions.

It is evident that the suggested estimator for the slope coefficients utilizes information from cross-sections in a similar way that the panel-based estimators adopted by Groen (2000) and Mark and Sul (2001) do. Like Stein, we are concerned with the estimation of several unknown coefficients. It is natural that these Stein-like estimator makes best use of existing information available. The implicit assumption underlying the use of information from cross-sections for our estimator, however, is very much different from that for the panel-based estimators. The panel-based estimators are built on the assumption that the slope coefficients are all the same for all the cross-sectional countries. The panel approach to pooling the data produces a single estimate. On the other hand, the shrinkage estimator allows for separate slope estimate for each cross-section country as the OLS estimator does, but makes use of the cross-sectional information that the OLS estimator does not. Thus both the shrinkage

estimator and the pooled estimator are the same to reduce the estimation errors, but differ in the way how the cross-sectional information is processed. Yet, our shrinkage estimator has the advantages of producing more reasonable slope estimates. When the truth is that the slope coefficients are all the same across each cross-sectional countries, the difference between the slope estimates from the two approaches should be insignificant, because by definition the shrinakge estimator embraces the extreme situation. But if the slope parameters are all different in each country, the panel approach by no means gives consistent slope estimate, because of a false constraint imposed in estimation, while our shrinkage estimator still does, as will be shown below.

The construction of the shrinkage estimator opens a few questions to be answered.

3.3 Shrinkage target

What is the rationale behind the choice of the grand average as the shrinkage target? The grand average has been commonly employed in the literature. Lindley (1962) demonstrates that the risk dominance of the Stein estimator when shrinking the sample mean toward the grand mean estimator. Our choice reflects a belief that the slope coefficients is different and near to each other. Like the constant target in the case of sample mean estimation, the grand average plays the role of a restricted estimator, while the OLS estimator is the unrestricted counterpart. Here the constraint imposed into our estimator is the slope coefficients of all the country are the same. Conventional econometric wisdom would suggest that imposing any structure or restriction in estimation brings forth efficiency gains, whether or not it is true. Consequently, our shrinkage estimator performs best when each of the slope coefficients coincides.

Alternative shrinkage targets as structures or constraints might exist. One shrinkage target works well for a certain data set, but not necessarily for another. It is thus difficult to tell which one works well a priori, without checking out-of-sample performance. This study looks at a dataset of aggregate time series from major developed countries. Within the circle, the transmission of shocks to fundamentals to exchange rates, if any, may vary with the macro environment or the aggregate policy in each country, but at a similar speed. The assumption of the slope heterogeneity is considered a reasonable one. We will show the advantages of the chosen target through simulations.

3.4 Shrinking factor

Applications of the shrinkage estimator is not possible without knowing the shrinking factor ω_i. The subsection is devoted to how to determine and estimate it. The major goal for all the existing estimators is in finite samples to minimize the estimation risk of the proposed estimators. This entails a need to specify a loss function. We propose to use the quadratic loss function. The loss function considers the tradeoff between bias and estimation errors using a quadratic distance measure between the true and the estimated slope coefficient. It gives rise to a risk function that calculates mean squared errors (MSE) as follows:

$$\Re(\tilde{\beta}_i, \beta_i) = \text{MSE}(\tilde{\beta}_i) = \text{E}\left[(\tilde{\beta}_i - \beta_i)^2\right].$$

Minimizing $\Re(\tilde{\beta}_i, \beta_i)$ with respect to ω_i gives

$$\omega_i = \frac{\text{MSE}(\bar{\beta}) - E\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right]}{\text{MSE}(\hat{\beta}_i) + \text{MSE}(\bar{\beta}) - 2E\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right]}.$$

We can describe the 'optimal' shrinking factor further, provided that the probability structure between the estimators is given. Without assuming normality, suppose the OLS estimator and the grand average in finite samples is jointly distributed as:

$$\begin{bmatrix} \hat{\beta}_i \\ \bar{\beta} \end{bmatrix} \sim \text{P}\left[\begin{pmatrix} \beta_i + \gamma_{i,ls} \\ \beta_i + \gamma_{i,g} \end{pmatrix}, \begin{pmatrix} \sigma_{i,ls}^2 & \rho_i \\ \rho_i & \sigma_{i,g}^2 \end{pmatrix}\right],$$

where P can be any distribution function as long as the first and second moments exist and are finite, and , and β_i is the true value of the slope coefficient i. Note that $\gamma_{i,ls}$ is the bias associated with the OLS estimator due to the high degree of persistence in predictors, and $\gamma_{i,g} = \frac{\sum_i \beta_i}{N} - \beta_i$, the distance of the true grand average from the true slope coefficient i. Therefore, when the slope coefficients are all the same, $\gamma_{i,g} = 0$. A important feature from the covariance matrix is to have the considered estimators correlated by an introduction of ρ_i. The appearance of the correlation measure is unusual when understanding the shrinkage estimator in an empirical Bayesian vein where it combines a prior, given by the grand average estimate, with the sample information corresponding to the least squares estimate. The typical empirical Bayesian approach assumes independence between prior and sample information, regardless of whether the prior is estimated from the same dataset as the grand average. Allowing for a correlation between the combined estimators may avoid damage in

estimations that takes place without considering the link.[2] With some algebra calculations,[3]

$$\omega_i = 1 - \frac{\sigma_{i,ls} + \gamma_{i,ls}^2 - \rho_i - \gamma_{i,ls}\gamma_{i,g}}{\sigma_{i,ls}^2 + \gamma_{i,ls}^2 + \sigma_{i,g}^2 + \gamma_{i,g}^2 - 2\rho_i - 2\gamma_{i,ls}\gamma_{i,g}}.$$

The expression offers some intuitions of how the weight is given to the combined estimators. Higher weight is assigned to the grand average when the OLS estimator has a higher bias and variance. But the estimator is drawn toward the OLS estimator in the case of high correlation between the combined estimators. When there are increases in $\gamma_{i,g}$ or $\sigma_{i,g}^2$, indications that the true slope coefficients are so scattered not as the prior prescribes, the OLS estimator receives more weight. This is because gains from information combination in both cases is small. Thus, very importantly, the weight adapts itself to the data.

It is crucial both in theory and in applications to estimate the shrinking factor consistently. A pre-requisite for establishing the consistency of the proposed estimator is the weight consistency.[4] Further, it controls the information quality represented by the combined estimators. A correct inference from the data is very much dependent on a precisely estimated weight. A technical innovation in this paper is to suggest a consistent bootstrap procedure to estimate the weight parameter. The available estimator as suggested by Judge and Mittelhammer (2004), though consistent, relies on the asymptotic argument, and its finite-sample performance can be satisfactory only when the sample size is sufficiently large. This might not be the case in the study because of a panel of short-spanned time series is under investigation. Typically the bootstrap estimator have proven to display more robust performance than the asymptotic counterpart. A more important justification for such a bootstrap procedure is that no general consistent estimator has been mounted for the bias parameters such as $\gamma_{it,l}$ and $\gamma_{i,g}$. We appeal to the bootstrap procedure in order to render our estimation and inference feasible in the subsequent empirical analysis. Very detailed illustration of the bootstrap algorithm is given in the appendix to save the space.

---

[2] One more difference is that the Bayesian approach is in general not optimized with any particular loss function.

[3] A detailed derivation is given in the appendix.

[4] We have shown in the appendix the consistency of the shrinkage estimator, given the existence of a consistent estimator for the shrinking factor.

4 Risk reduction: simulation analysis

This section investigates to what extent the the proposed estimator can improve over the traditional estimator in terms of risk reduction through simulations. To make our analysis more realistic, the DGP calibrated to the actual data is considered. Further, to permit a direct comparison, the data under simulation analysis now and empirical study later is the same as that used in Kilian (1999).[5] The data spans from 1973.I to 1997.IV, and contains the US dollar exchange rates of the Canadian dollar, the German mark, the Japanese yen, and the Swiss franc. Thus, the country number $N = 4$ in the study.

The DGPs are based on both the restricted VEC model under the null that the exchange rate is unpredictable, and under the alternative that the fundamental helps to predict the exchange rate, on the maintained assumption that the exchange rate and the fundamental are cointegrated. The key difference in generating data under the null or the alternative thus lies in whether β_i is set to 0. The lag orders of $\triangle x_{it}$ is determined using the Campbell-Perron's general-to-specific rule. A seemingly uncorrelated regression is then fit to the data after the lag orders are settled. Each country is estimated a separate DGP, while the residuals are allowed to be dependent cross-sectionally measured by the joint error covariance matrix Σ. Vector sequence of the innovations are now drawn from $N(0, \hat{\Sigma})$, where the covariance matrix is estimated from the SUR residuals. The pseudo observations on $\triangle s_{it}$ and $\triangle x_{it}$ are produced recursively based on the estimated version of (6) and (7). The MSEs under the null and the alternative are computed and recorded individually using the artificial data from 2000 replications.

Table (1) reports the relative empirical risks of the shrinkage estimator and the least-square estimator under the null and the alternative processes, respectively. The important and encouraging message coming out of the simulations is that the shrinkage estimator empirically dominates the least-square estimator, regardless of which simulation scenario is considered. But the dominance weakens slightly as the forecast horizon increases. Virtually the risk reductions using the shrinkage estimator can be as large as between 10% and 35%. A decomposition of the MSE clearly reveals that the risk reductions using the shrinkage estimator attributes much more to the reductions in the estimation errors, implying accuracy

---

[5]All the data is available and can be downloaded from the homepage of *Journal of Applied Econometrics.*

gains. In some instances, the bias even increases due to use of the shrinkage estimator, although the bias increases do not offset the error reductions.

The risk improvements of the shrinkage estimator is embodied further into power gains in testing. Table (2) presents the size-adjusted empirical power of using the shrinkage estimator as the test statistic against the predictability alternative. An inspection of Table (2) shows that the power gains from using the shrinkage estimator is 10% or more in many cases. For Germany data, the power gain can be even up to 20%. An implication of the finding is merely that the predictability alternative can now be better detected from the data when the test statistics are based on the shrinkage estimator. Plot (1) depicts the clear difference between the finite-sample distributions of the shrinkage estimator and the least-square estimator, based on the Germany data. [6] In which, the distributions of the shrinkage estimator under the null and the alternative are more centered. This is again due to the error reductions, and exactly where power gains come.

5   A re-examination of the exchange rate predictability

We are in a position to re-examine the empirical validity of the exchange rate predictability applying the shrinkage estimator. The testing strategy basically follows that utilized in Mark (1995), Kilian (1999), and Mark and Sul (2001). These studies all base their inference on the bootstrap approach in order to control for small-sample bias for which the asymptotic approximation generally fails to correct. There additional considerations for use of the estimator. As discussed, the estimate for the optimal weight is given based on the bootstrap. Next, the asymptotic theory for the proposed estimator has not been established yet. It is expected that the asymptotic distributions of the statistics based on the shrinkage estimator is ill approximated by the normal, because of the high degree of persistence characteristic in predictors.

Before the asymptotic theory for the estimator is derived, inference on the predictability here is mainly drawn from the bootstrap. The bootstrap conducted is very much the same as described in the preceding simulation section, except the step about sampling innovations. Instead of drawing samples from the parametric normal distribution, the re-samples are now

---

[6]Similar distribution patterns are also uncovered from other country's data. They are not reported here to economize the presentation.

drawn from a sequence of the restricted residual vectors. In other words, a non-parametric bootstrap is employed. The restriction again is no exchange rate predictability, equivalently $\beta_i = 0$.

A by-product of the study is to test whether the exchange rate is cointegrated with the fundamental, based on the shrinkage estimator. This can be done simply by imposing another restriction that $\alpha_i = 0$ in (7) in the aforementioned re-sampling schemes.

5.1   A new look at cointegration

As has been emphasized, when the cointegration between the exchange rate and the fundamental does not come into existence, the exchange rate is simply unpredictable. We report the results for testing for cointegration using the shrinkage estimator in Table (3).

The bootstrap coefficient test easily rejects the null of no cointegration between the exchange rate and the fundamental, mostly at 5% significance level, with the exception of Japan with $k = 12$. The testing results based on R^2 lend additional credibility to the notion of a cointegration between the exchange rate and the fundamental. Most of the corresponding p-values for R^2 are smaller than 5%, remarkably similar to those for the coefficient test.

Some obvious patterns on the estimated coefficients can be summarized from table. The estimates for the slope coefficients are all shrunk further toward the grand average for all the countries and all the forecast horizons (k). In general, the shrinkage estimates are less dispersed than the OLS counterparts. This is to be expected from the shrinkage theorem. For any particular country, the estimated slope coefficients increases as the horizons increases. Nevertheless this should not be interpreted as an evidence for the increased long-horizon predictability, as the cointegration can not be interchanged with the predictability in notions. More importantly, the estimated slope coefficients for any fixed horizon are clearly seen to vary to a certain degree. This appears to be consistent with our prior belief, but contradicts the extreme assumption made in the panel-based approach that the slope coefficients are the same for all the countries.

We now turn to the estimated optimal shrinking factors. The homogeneity in slope coefficients would suggest that the estimated weights are similar in magnitude for any k. A close inspection of the estimates shows however little evidence for the slope homogeneity. In particular, some of the weights for Germany and Japan are negatively estimated. It is due to

that no constraint on the weight to be positive and smaller than one was placed in estimation. Taking the normality approximation, if not too arbitrary, these negative weight estimates are all covered in confidence intervals centered around the origin at 95% level, implying a true weight indifferent from zero. In other words, the grand averages receives most of the weight. Typically this might take place when the least square estimates are not reliable for higher variance or larger bias. These are also the cases where the risk improvements over the least square estimator may be substantial.

5.2   Out-of-sample forecast adjusted for risks

We access the relative forecast accuracy of the two competing models with DM (Diebold and Mariano, 1995) statistic and Theil's U statistic. It should be noted that the problem with estimating the long-run variance precisely often leads to spurious inference, as documented in the literature.

The forecast results are displayed in Tables (4) and (5). The former is based on the DM statistic, and the latter the Theil's U statistic. There is now stronger evidence presented for the dominance of the monetary model over the random walk, after accounting for estimation risks using the shrinkage estimates. With few exceptions, the p-values associated with the shrinkage estimates for both statistics are smaller, relative to those associated with the least square estimates.

It stands out from the results that controlling over the risks uncovers more favorable evidence in supports of the monetary model, while there is essentially no evidence for so when leaving the risks unattended. Instances of this are found more in Table (5) than in Table (4). For example, at almost all horizons, the monetary model is found to be superior to the random walk in terms of predictability for Germany and Japan. This contrasts sharply with the previous findings where no predictability is reported. Considering the Theil's U statistic is more robust, this evidence lends quite a good deal of credence to the predictability at long-horizon. Besides, there seems to have evidence of declining p-values as the prediction horizon is increased, especially the results based on the Theil's U statistic for all the countries but Canada. Overall, by adjusting for the estimation risks, the evidence that the fundamentals help forecast the exchange rate at short or even long horizon is stronger and clearer.

6   Conclusion

It is no denying that inference on the exchange rate predictability based on the univariate predictive regression is not trustworthy due to estimation risks. The panel-based approach exploiting cross-sectional information may reduce estimation errors, but incurs another type of estimation risk with the extreme assumption of the parameter homogeneity. The shrinkage approach proposed enjoys risk gains, while allowing for the parameter heterogeneity. Our Monte-Carlo simulations further demonstrate that the merits of applying the shrinkage approach converts into power gains in the testing context. Significant evidence for the exchange rate predictability can now be found at least for short-horizons. Whether there is evidence of higher predictability at longer horizons remains unconfirmed, although more evidence points to be so after controlling over estimation risks.

One interesting question that deserves more research efforts is to investigate the effects of the country number on the robustness of the inference. The shrinkage approach is basically concerned with estimations of many unknown parameters. This is suggestive of further risk improvements from utilizing more cross-sectional information. Rigorous research might address to what degree an increase in country number can help reduce risks, and how much economic value of additional risk reductions is worth.

Appendix 1: Derivation of the shrinking factor

The risk function as defined based on our shrinkage estimator can be expressed as

$$\text{MSE}(\tilde{\beta}(\omega_i)) = \omega_i^2 \text{MSE}(\hat{\beta}_i) + (1 - \omega_i)^2 \text{MSE}(\bar{\beta}) + 2\omega_i(1 - \omega_i)\text{E}\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right].$$

Minimizing it with respect to ω_i gives the first-order condition as

$$2\omega_i \text{MSE}(\hat{\beta}_i) - (2 - 2\omega_i)\text{MSE}(\bar{\beta}) + (2 - 4\omega_i)\text{E}\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right] = 0,$$

from which the optimal shrinking factor can be found

$$\omega_i = \frac{\text{MSE}(\bar{\beta}) - \text{E}\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right]}{\text{MSE}(\hat{\beta}_i) + \text{MSE}(\bar{\beta}) - 2\text{E}\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right]}.$$

as described. Note that

$$\mathrm{MSE}(\hat{\beta}_i) = \sigma_{i,ls}^2 + \gamma_{i,ls}^2,$$

$$\mathrm{MSE}(\bar{\beta}) = \sigma_{i,g}^2 + \gamma_{i,g}^2,$$

$$\mathrm{E}\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right] = \rho_i + \gamma_{i,ls}\gamma_{i,g}.$$

Replacing with all the corresponding expressions above in the equation concerning the shrinking factor then complete the derivation.

Moreover, it is easy to show that the optimal weight derived satisfies the second-order condition, because

$$\mathrm{MSE}(\hat{\beta}_i) + \mathrm{MSE}(\bar{\beta}) - 2 \cdot \mathrm{E}\left[(\hat{\beta}_i - \beta_i)(\bar{\beta} - \beta_i)\right] = \mathrm{E}\left[(\hat{\beta}_i - \bar{\beta})^2\right] > 0.$$

In other words, $\tilde{\beta}_i$ is the minimum among all linear combinations of $\hat{\beta}_i$ and $\bar{\beta}$ under the MSE criterion.

Appendix 2: The consistency of the shrinkage estimator

We provide some regularity conditions under which the consistency can be established. Suppose that $T\hat{\sigma}_{i,ls}^2 \overset{p}{\to} \sigma_{i,ls}^2$, $T\hat{\sigma}_{i,g}^2 \overset{p}{\to} \sigma_{i,g}^2$, and $T\hat{\rho}_i \overset{p}{\to} \rho_i$. Also assume that $\hat{\gamma}_{i,ls} - \gamma_{i,ls} \equiv o_p(T^{-1})$, $\hat{\gamma}_{i,g} - \gamma_{i,g} \equiv O_p(T^{-1})$ and $\hat{\beta}_i - \beta_i \equiv O_p(\sqrt{T}^{-1})$. Then, when $\sigma_{i,g}^2 \neq 0$,

$$\mathrm{plim}(b\tilde{e}ta(\hat{\omega}_i)) = \mathrm{plim}\left[\hat{\omega}_i\hat{\beta}_i + (1 - \hat{\omega}_i)\bar{\beta}\right]$$

$$=\mathrm{plim}(\hat{\beta}_i) + \mathrm{plim}\left[\frac{\hat{\sigma_{i,ls}}^2 + \hat{\gamma_{i,ls}}^2 - \hat{\rho}_i - \hat{\gamma}_{i,ls}\hat{\gamma}_{i,g}}{(\hat{\beta}_i - \bar{\beta})^2} \cdot (\hat{\beta}_i - \bar{\beta})\right]$$

$$=\beta_i + \frac{O_p(T^{-1}) + o_p(T^{-1}) \cdot o_p(T^{-1}) - O_p(T^{-1}) - o_p(T^{-1}) \cdot O_p(T^{-1})}{O_p(1) \cdot O_p(1)} \cdot O_p(1)$$

$$=\beta_i + o_p(1) = \beta_i.$$

But if $\sigma_{i,g}^2 = 0$, $\hat{\beta}_i - \beta_i \equiv O_p(\sqrt{T}^{-1})$,

$$\mathrm{plim}(\tilde{\beta}(\hat{\omega}_i))$$

$$=\mathrm{plim}(\hat{\beta}_i) + \mathrm{plim}\left[\frac{T\hat{\sigma_{i,ls}}^2 + T\hat{\gamma_{i,ls}}^2 - T\hat{\rho}_i - T\hat{\gamma}_{i,ls}\hat{\gamma}_{i,g}}{T(\hat{\beta}_i - \bar{\beta})^2} \cdot (\hat{\beta}_i - \bar{\beta})\right]$$

$$=\beta_i + \frac{O_p(1) + o_p(1) \cdot o_p(T^{-1}) - O_p(1) - o_p(1) \cdot o_p(T^{-1})}{O_p(1) \cdot O_p(1)} \cdot o_p(1)$$

$$=\beta_i + O_p(1) \cdot o_p(1) = \beta_i.$$

This proves the consistency for the estimator.

Appendix 3: Bootstrap algorithm to estimating the optimal weight

The bootstrap procedures are to obtain robust estimates of the parameters appearing in the optimal shrinking factor. The detailed illustrations are as follows:

1. Estimate $\triangle e_{it+k} = \beta_i x_{it} + \varepsilon_{it+k}$ and $\varepsilon_{i,t+k} = \sum_{l=1}^{m} \rho_{il}\varepsilon_{i,t+k-l} + v_{i,t+k}$ with OLS, in which the lag order m is chosen by the AIC criteria, and obtain $\hat{\beta}_i$, $\hat{\rho}_{il}$, $\hat{\varepsilon}_{i,t+k}$ and $\hat{v}_{i,t+k}$.

2. Now draw from $\{\hat{v}_{i,t+k}\}$ with replacements, and generate re-samples using the data-generating process: $\triangle e^*_{it+k} = \hat{\beta}_i x_{it} + \hat{\varepsilon}^*_{it+k}$ and $\hat{\varepsilon}^*_{i,t+k} = \sum_{l=1}^{m} \hat{\rho}_{il}\hat{\varepsilon}^*_{i,t+k-l} + \hat{v}_{i,t+k}$.

3. Regress $\triangle \hat{e}^*_{it+k}$ against x_{it} with OLS and obtain the bootstrap estimates of the slope coefficient, denoted by $\hat{\beta}^*_i$.

4. Repeat steps 2 and 3 B_0 times. Now compute the average bias by $\frac{\sum_{b=1}^{B_0} \hat{\beta}_i - \hat{\beta}^*_{i,b}}{B_0}$, and deduct it from $\hat{\beta}_i$. This is the bootstrap bias-corrected estimate for the slope coefficient, labeled as $\hat{\hat{\beta}}_i$. Then the biased-corrected residuals can be computed accordingly by $\hat{\hat{\varepsilon}}_{it+k} = \triangle e_{it+k} - \hat{\hat{\beta}}_i x_{it}$.

5. Now generate the bootstrap samples based on $\hat{\hat{\beta}}_i$, $\hat{\hat{\varepsilon}}_{i,t+k}$, x_i as if they are true ones by repeating B_0 times steps 2, 3, and 4. Thus, a sequence of $\{\hat{\hat{\beta}}^*_i\}$ is generated. Repeat the same procedures for different countries other than i, and obtain the bootstrap grand average sequence of $\{\bar{\bar{\beta}}^*\}$ for the N countries.

Based on the two sequences of $\{\hat{\hat{\beta}}^*_i\}$ and $\{\bar{\bar{\beta}}^*\}$ obtained from the aforementioned procedures, the bootstrap estimate for the derived optimal shrinkage factor is computed as

$$\hat{\omega}^*_i = \frac{\sum(\bar{\bar{\beta}}^* - \hat{\hat{\beta}}_i)^2 - \sum(\hat{\hat{\beta}}^*_i - \hat{\hat{\beta}}_i)(\bar{\bar{\beta}}^* - \hat{\hat{\beta}}_i)}{\sum(\hat{\hat{\beta}}^*_i - \hat{\hat{\beta}}_i)^2 + \sum(\bar{\bar{\beta}}^* - \hat{\hat{\beta}}_i)^2 - 2\sum(\hat{\hat{\beta}}^*_i - \hat{\hat{\beta}}_i)(\bar{\bar{\beta}}^* - \hat{\hat{\beta}}_i)}.$$

References

[1] Berben, R.B. and D.J. van Dijk (1998), "Does the Absence of Cointegration Explain the Typical Findings in Long Horizon Regressions," Papers 9814/a, Erasmus University of Rotterdam − Econometric Institute.

[2] Berkowitz, J. and L. Giorgianni (2001), "Long-Horizon Exchange Rate Predictability?" *Reviews of Economics and Statistics*, 83(1), 81-91.

[3] Boswijk, H.P. (1994), "Testing for an Unstable Root in Conditional and Structural Error Correction Models," *Journal of Econometrics*, 63(1), 37-60.

[4] Cheung, Y-W, M.D. Chin, and A. G. Pascual (2005), "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?" *Journal of International Money and Finance*, 24, 1150−1175.

[5] Chinn, M.D. and R.A. Meese (1995), "Banking on Currency Forecasts: How Predictable is Change in Money?" *Journal of International Economics*, 161−178.

[6] Diebold, F.X. and L. Kilian (2000), "Unit Root Tests are Useful for Selecting Forecasting Models," *Journal of Business and Economic Statistics*, 18, 265−273.

[7] Dumas, B. and B. Jacquillat (1990), "Performance of Currency Portfolios Chosen by a Baysian Technique: 1967−1985," *Journal of Banking and Finance*, 14, 539−558.

[8] Engel, C. and K.D. West (2005), "Exchange Rate and Fundamentals," *Journal of Political Economy*, 113, 485−517.

[9] Engle, R.F., D.F. Hendry, and J.-F. Richard (1983), "Exogeneity,", *Econometrica*, 51, 277−307.

[10] Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, 7(1), 1-26.

[11] Granger, C. W. J. and Newbold, P. (1974), "Spurious regressions in econometrics," *Journal of Econometrics*, 2, 111-120.

[12] Greene, W.H. (2000), *Econometric Analysis*, Prentice Hall, 4th Edition.

[13] Groen, J.J.J. (1999), "Long Horizon Predictability of Exchange Rates: Is it for Real?" *Empirical Economics*, 24, 451−469.

[14] Groen, J.J.J. (2000), "The Monetary Exchange Rate Model as a Long-Run Phenomenon," *Journal of International Economics*, 52, 299−319.

18

[15] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.

[16] Hansen, L.P. and R.J. Hodrick (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, 829−853.

[17] James, W. and C.M. Stein (1961), "Estimation with Quadratic Loss," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability (vol. 1)*, Berkeley, California: University of California Press, 361−380.

[18] Jobson, J.D., B. Korkie and V. Ratti (1979),"Improved Estimation for Markowitz Portfolios Using James-Stein Type Estimators," *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, 41, 279−284.

[19] Jorion, P. (1985), "International Portfolio Diversification with Estimation Risk," *Journal of Business*, 58(3), 259−278.

[20] Jorion, P. (1986), "Bayes-Stein Estimation for Portfolio Analysis," *Journal of Financial and Quantitative Analysis*, 21, 279−291.

[21] Jorion, P. (1991), "Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection," *Journal of Banking and Finance*, 10, 717−727.

[22] Judge, G. G., W. E. Griffiths, R. C. Hill, and T.C. Lee (1980), *The Theory and Practice of Econometrics*, New York: John Wiley & Sons.

[23] Judge, G.G. and M.E. Bock (1978), *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*, Amsterdam: North−Holland.

[24] Judge G. G. and R. Mittelhammer (2004), "A Semiparametric Basis for Combing Estimation Problems under Quadratic Loss," *Journal of American Statistical Association*, 99, 479−487.

[25] Kendall, M.G. (1954), "Note on the Bias in the Estimation of Autocorrelation," *Biometrika*, 41, 403-404.

[26] Kilian L. (1999),"Exchange Rates and Monetary Fundamentals: What Do We Learn from Long Horizon Regressions?" *Journal of Applied Econometrics*, 14, 491−510.

[27] Lindley, D.V. (1962), "Discussion of Professor Stein's Paper," *Journal of the Royal Statistical Society*, Series B, 24, 285-288.

[28] MacKinnon, J.G. and Smith A.A. (1998), "Approximate Bias Correction in Econometrics," *Journal of Econometrics*, 85, 205−230.

[29] Mark, N.C. (1995),"Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85, 201-218.

[30] Mark, N.C. and D. Sul (2001), "Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Seventeen Country Panel," *Journal of International Economics*, 53, 29−52.

[31] Marriott, F.H.C. and J.A. Pope (1954), "Bias in the Estimation of Autocorrelations," *Biometrika*, 41, 393-402.

[32] Meese, R.A. and K. Rogoff (1983), "Emprircal Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*, 14, 3-74.

[33] Rossi, B. (2005), "Testing Long-Horizon Predictive Ability with High Persistence, and the Meese-Rogoff Puzzle," *International Economic Review*, 46, 61−92.

[34] Sclove, S.L., C. Morris, and R. Radhakrishman (1972), "Non Optimality of Preliminary Test Estimators for the Multinormal Mean," *Annals of Mathematical Statistics*, 43, 1481-1490.

[35] Stambaugh, R.F. (1999), "Predictive Regressions," *Journal of Financial Economics*, 54, 375−421.

[36] Stein, C.M. (1955), "Inadmissibility of the Mean of a Multivariate Normal Distribution," in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability (vol. 1)*, Berkeley, California: University of California Press, 197−206.

[37] West, K.D. (1996), "Asymptotic Inference about Predictive Ability," *Econometrica*, 64(5), 1067−1084.

[38] Zellner, A. and W. Vandaele (1974), "Bayes-Stein Estimators for $k-$Means, Regression and Simultaneous Equation Models," in S.E. Fienberg and A. Zellner, eds., *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, North$-$Holland, Amsterdam, 627-653.

[39] Zviot, E. (1996), "The Power of Single Equation Tests for Cointegration when the Cointegrating Vector is Prespecified," *Working paper*, Department of Economics, University of Washington.
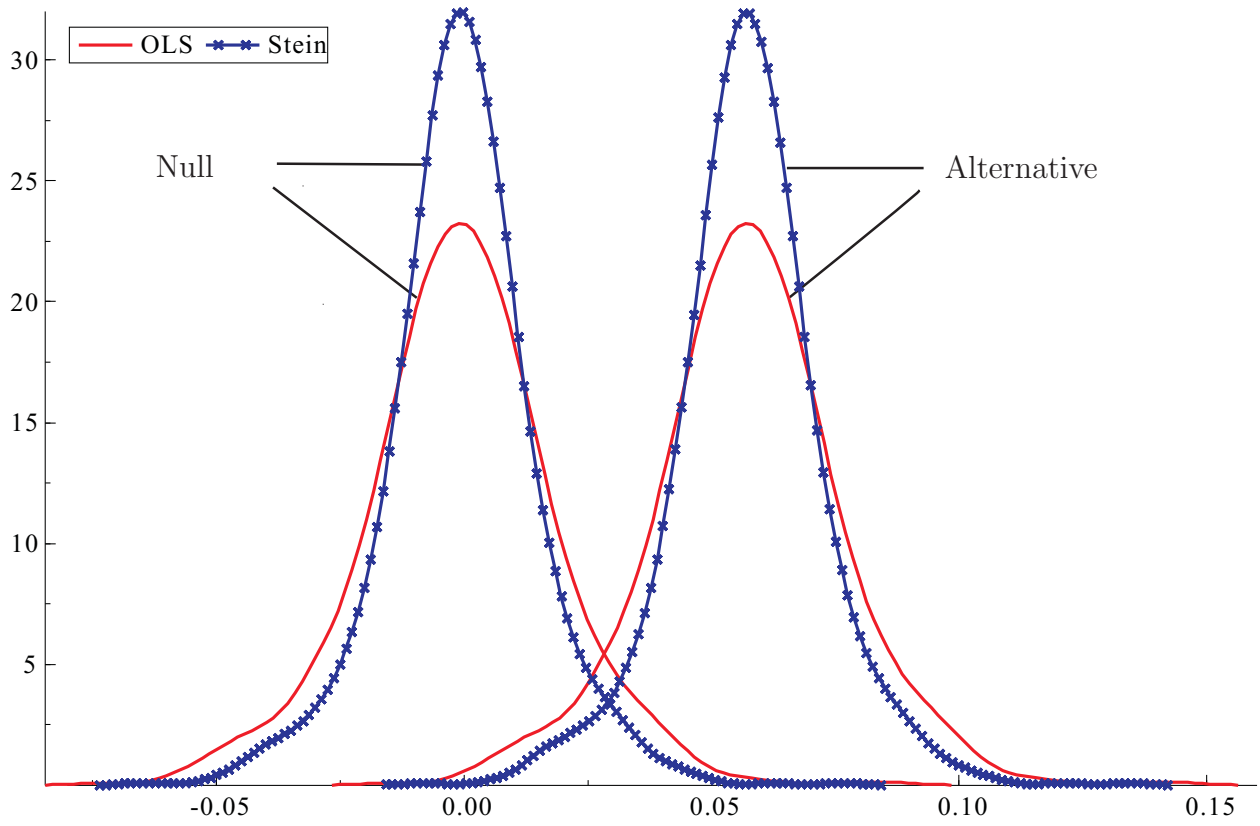
Figure 1: The small-sample distributions of the shrinkage and OLS Estimators under H_0 and H_a

Table 1: Relative estimation risk (Shrinkage/LS)

Country/Horizons	1		4		8		12		16	
	H_0 [b]	H_a [c]	H_0	H_a	H_0	H_a	H_0	H_a	H_0	H_a
MSE: [a]										
Canada	0.829	0.830	0.817	0.822	0.801	0.836	0.817	0.857	0.848	0.865
Germany	0.667	0.669	0.718	0.722	0.672	0.754	0.687	0.774	0.695	0.827
Japan	0.621	0.624	0.701	0.774	0.655	0.809	0.682	0.840	0.706	0.861
Switzerland	0.800	0.840	0.822	0.851	0.806	0.858	0.816	0.862	0.852	0.892
bias^2:										
Canada	0.875	0.763	0.422	1.209	0.912	1.502	0.733	1.423	1.236	1.201
Germany	0.626	0.891	0.963	1.104	0.802	0.825	0.832	0.861	1.153	1.424
Japan	1.632	1.230	0.870	1.429	0.637	1.410	0.816	1.520	0.822	1.535
Switzerland	1.422	1.229	1.352	1.341	0.403	0.854	0.534	0.732	0.833	0.778
variance:										
Canada	0.830	0.830	0.818	0.813	0.802	0.784	0.817	0.805	0.848	0.818
Germany	0.667	0.669	0.718	0.720	0.672	0.768	0.687	0.785	0.695	0.805
Japan	0.621	0.624	0.701	0.707	0.655	0.749	0.683	0.779	0.707	0.802
Switzerland	0.800	0.837	0.822	0.841	0.846	0.864	0.876	0.874	0.893	0.903

[a] MSE is defined as the sum of the bias squared and the variance of the parameter estimates:

$$\text{MSE}(\tilde{\beta}_i) = \text{E}\left[(\tilde{\beta}_i - \beta_i)^2\right] = \text{E}\left\{\left[(\tilde{\beta}_i - \text{E}(\tilde{\beta}_i)) + (\text{E}(\tilde{\beta}_i) - \beta_i)\right]^2\right\} = \text{variance}(\tilde{\beta}_i) + \text{bias}(\tilde{\beta}_i)^2.$$

[b] The column gives the ratios of the MSE, bias squared, and variance of the Shrinkage estimates to by those of the OLS counterparts, under the null of no predictability.

[c] The column gives the ratios of the MSE, bias squared, and variance of the Shrinkage estimates to by those of the OLS counterparts, under the alternative of predictability.

Table 2: Power performance of the shrinkage estimator

(a) 5% significance level

Country/Horizons	1	4	8	12	16
Canada	0.845[a]	0.850	0.854	0.792	0.723
	(1.095)[b]	(1.108)	(1.131)	(1.145)	(1.172)
Germany	0.616	0.680	0.607	0.540	0.507
	(1.279)	(1.444)	(1.317)	(1.239)	(1.213)
Japan	0.826	0.778	0.863	0.807	0.760
	(1.309)	(1.319)	(1.214)	(1.171)	(1.193)
Switzerland	0.735	0.715	0.757	0.720	0.628
	(1.030)	(1.033)	(1.080)	(1.136)	(1.150)

(b) 10% significance level

Country/Horizons	1	4	8	12	16
Canada	0.925	0.941	0.925	0.868	0.790
	(1.085)	(1.079)	(1.087)	(1.081)	(1.078)
Germany	0.802	0.783	0.804	0.770	0.713
	(1.303)	(1.288)	(1.247)	(1.207)	(1.135)
Japan	0.912	0.896	0.920	0.907	0.847
	(1.078)	(1.106)	(1.108)	(1.106)	(1.107)
Switzerland	0.896	0.849	0.867	0.818	0.754
	(1.083)	(1.060)	(1.074)	(1.068)	(1.094)

[a] The entries represent size-adjusted power of the Shrinkage estimator against the alternative of exchange rate predictability.

[b] The entries in parentheses represent the power performance of the Shrinkage estimator relative to that of the OLS estimator (Shrinkage/OLS).

Table 3: Full-sample estimation results and tests for cointegration

statistics	$\tilde{\beta}_i$[a]	p-value[b]	R^2	p-value[c]	$\hat{\beta}_i$	$\bar{\beta}$	$\hat{\omega}_i$	$\text{std}(\hat{\omega}_i)$[d]
Canada:								
1	0.035	**0.000**	0.035	**0.019**	0.029	0.052	0.772	0.311
4	0.131	**0.000**	0.104	**0.010**	0.106	0.207	0.734	0.285
8	0.285	**0.000**	0.218	**0.011**	0.237	0.428	0.727	0.289
12	0.370	**0.011**	0.194	**0.031**	0.230	0.625	0.742	0.314
16	0.379	**0.012**	0.133	**0.070**	0.291	0.790	0.813	0.342
Germany:								
1	0.051	**0.000**	0.036	**0.032**	0.045	0.052	0.125	0.340
4	0.195	**0.021**	0.119	**0.004**	0.178	0.207	0.183	0.295
8	0.412	**0.037**	0.214	**0.018**	0.385	0.428	-0.035	0.287
12	0.599	**0.042**	0.340	**0.013**	0.617	0.625	-0.047	0.320
16	0.767	**0.004**	0.483	**0.009**	0.832	0.790	0.083	0.346
Japan:								
1	0.051	**0.000**	0.037	**0.027**	0.049	0.052	0.100	0.328
4	0.199	**0.042**	0.121	**0.006**	0.207	0.207	0.008	0.301
8	0.408	**0.032**	0.229	**0.010**	0.454	0.428	-0.067	0.310
12	0.586	**0.062**	0.311	**0.012**	0.717	0.625	-0.111	0.324
16	0.734	**0.006**	0.376	**0.019**	0.947	0.790	-0.145	0.347
Switzerland:								
1	0.072	**0.000**	0.064	**0.000**	0.087	0.052	0.578	0.305
4	0.275	**0.000**	0.221	**0.000**	0.336	0.207	0.557	0.321
8	0.542	**0.000**	0.366	**0.001**	0.634	0.428	0.589	0.301
12	0.787	**0.000**	0.520	**0.000**	0.874	0.625	0.684	0.305
16	1.048	**0.001**	0.722	**0.000**	1.090	0.790	0.874	0.320

[a] Shrinkage estimates is defined as $\tilde{\beta}_i = \hat{\omega}_i \hat{\beta}_i + (1 - \hat{\omega}_i)\bar{\beta}$, where $\hat{\beta}_i$ is the OLS estimate of the slope and the $\bar{\beta}$ is the grand average of the OLS estimates of the slopes of the 4 countries; $\hat{\omega}_i$ is the estimated optimal weight.

[b,c] P-value under the null of no exchange rate predictability ($\tilde{\beta}_i=0$ and $R^2=0$, respectively). Bold-faced numbers refer to p-values less than 10%.

[d] Standard deviation of the optimal weight estimates ($\hat{\omega}_i$).

Table 4: Out-of-Sample forecast evaluations: DM Statistic

Country	k	DM(A)[a]	p-value	p-value^{Kc}	DM(20)[b]	p-value	p-value^{Kd}
Canada	1	1.568	**0.015**	**0.041**	5.719	**0.000**	**0.027**
	4	1.500	**0.027**	**0.057**	1.786	**0.030**	**0.048**
	8	1.269	**0.065**	**0.015**	1.269	**0.070**	**0.016**
	12	-0.518	0.333	**0.064**	-0.523	0.319	**0.070**
	16	-1.378	0.655	0.110	-1.311	0.603	0.117
	max[e]	1.568	**0.052**	**0.060**	5.719	**0.005**	**0.056**
Germany	1	0.534	**0.095**	0.151	0.841	**0.093**	0.141
	4	0.438	0.144	0.162	0.529	0.139	0.160
	8	0.433	0.169	0.218	0.433	0.180	0.216
	12	0.627	0.186	0.249	0.603	0.196	0.250
	16	0.796	0.191	0.321	0.777	0.195	0.314
	max	0.796	0.288	0.273	0.841	0.300	0.268
Japan	1	0.770	**0.080**	**0.082**	0.990	**0.094**	0.104
	4	0.614	0.134	0.151	0.745	0.131	0.144
	8	0.855	0.129	0.133	0.890	0.136	0.133
	12	0.772	0.161	0.270	0.747	0.180	0.269
	16	0.729	0.170	0.451	0.717	0.180	0.433
	max	0.855	0.283	0.240	0.990	0.280	0.253
Germany	1	2.658	**0.003**	**0.010**	3.482	**0.003**	**0.023**
	4	2.586	**0.004**	**0.019**	2.718	**0.009**	**0.024**
	8	1.997	**0.033**	**0.045**	2.599	**0.018**	**0.039**
	12	1.744	**0.045**	**0.063**	1.981	**0.034**	**0.064**
	16	1.447	**0.063**	**0.077**	1.482	**0.080**	**0.097**
	max	2.658	**0.038**	**0.091**	3.482	**0.021**	**0.089**

Notes: The DM statistic is defined as $DM = \bar{d}/\sqrt{2\pi\hat{f}_d(0)/N_f}$, where $\bar{d} = N_f^{-1}\sum_{t=t_0+k}^{T}(u_{r,t}^2 - u_{m,t}^2)$ with $u_{r,t}$ and $u_{m,t}$ refer to the forecast errors of the random walk model and the monetary model, respectively. N_f is the number of recursive forecasts, and t_0 is the first date of forecast. $f_d(0)$ is the spectral density of $(u_{m,t}^2 - u_{r,t}^2)$ evaluated at frequency 0. Its consistent estimate $\hat{f}_d(0)$ is obtained using Newey and West (1987).

[a] DM statistic computed with truncation lags under Bartlett window set to 20.

[b] DM statistic with truncation lags under Bartlett window set by Andrews's (1991) algorithm.

[c,d] The corresponding p-value in Kilian (1999). Bold-faced numbers are those significant at 10% level.

[e] Joint test statistic proposed by Mark (1995), taking the maximum of a sequence of DM statistics indexed by k.

Table 5: Out-of-sample forecast evaluations: Theil's U

Country	k	Theil's U	p-value	p-value^{K}[a]
Canada	1	0.971	**0.009**	**0.042**
	4	0.936	**0.030**	**0.068**
	8	0.900	**0.043**	**0.057**
	12	1.063	0.549	**0.094**
	16	1.179	0.761	0.151
	min[b]	0.900	**0.081**	0.137
Germany	1	0.989	**0.061**	0.150
	4	0.979	**0.097**	0.169
	8	0.961	0.109	0.218
	12	0.864	**0.044**	0.291
	16	0.729	**0.010**	0.473
	min	0.729	**0.013**	0.283
Japan	1	0.987	**0.041**	0.118
	4	0.968	**0.090**	0.157
	8	0.927	**0.069**	0.145
	12	0.886	**0.051**	0.274
	16	0.838	**0.045**	0.534
	min	0.838	**0.053**	0.250
Swizerland	1	0.973	**0.006**	**0.017**
	4	0.929	**0.021**	**0.029**
	8	0.873	**0.025**	**0.045**
	12	0.808	**0.015**	**0.030**
	16	0.696	**0.000**	**0.026**
	min	0.696	**0.003**	**0.031**

Notes: Theil's U-statistic is defined as the ratio of the root-mean-square prediction error of the monetary model based on the shrinkage estimator to that of the random walk model. The null hypothesis is that the two models provide forecasts of equal accuracy (U=1). The alternative hypothesis is that the monetary fundamentals is more accurate (U<1).

[a] The corresponding p-values in Kilian (1999). Bold-faced numbers are those significant at 10% level.

[b] Joint test statistic according to Mark (1995), which takes the minimun of a sequence of Theil's U statistics indexed by k.