Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 963

January 2009

The Taylor Rule and Interval Forecast For Exchange Rates

Jian Wang and Jason J. Wu

The Taylor Rule and Interval Forecast For Exchange Rates

Jian Wang and Jason J. Wu*

Abstract: This paper attacks the Meese-Rogoff (exchange rate disconnect) puzzle from a different perspective: out-of-sample interval forecasting. Most studies in the literature focus on point forecasts. In this paper, we apply Robust Semi-parametric (RS) interval forecasting to a group of Taylor rule models. Forecast intervals for twelve OECD exchange rates are generated and modified tests of Giacomini and White (2006) are conducted to compare the performance of Taylor rule models and the random walk. Our contribution is twofold. First, we find that in general, Taylor rule models generate tighter forecast intervals than the random walk, given that their intervals cover out-of-sample exchange rate realizations equally well. This result is more pronounced at longer horizons. Our results suggest a connection between exchange rates and economic fundamentals: economic variables contain information useful in forecasting the distributions of exchange rates. The benchmark Taylor rule model is also found to perform better than the monetary and PPP models. Second, the inference framework proposed in this paper for forecast-interval evaluation, can be applied in a broader context, such as inflation forecasting, not just to the models and interval forecasting methods used in this paper.

Keywords: the Exchange Rate Disconnect Puzzle; Exchange Rate Forecast; Interval Forecasting.

JEL Codes: C14, C53, and F31

* Author notes: Wang is a senior economist in the Research Department of the Federal Reserve Bank of Dallas. He can be reached at Jian.Wang@dal.frb.org. Wu is an economist in the Division of Banking Supervision and Regulation of the Federal Reserve Board. He can be reached at Jason.J.Wu@frb.gov. We wish to thank Menzie Chinn, Charles Engel, Bruce Hansen, Mark Wynne, and Ken West for invaluable discussions. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

1 Introduction

Recent studies explore the role of monetary policy rules, such as Taylor rules, in exchange rate determination. They find empirical support in these models for the linkage between exchange rates and economic fundamentals. Our paper extends this literature from a different perspective: interval forecasting. We find that the Taylor rule models can outperform the random walk, especially at long horizons, in forecasting twelve OECD exchange rates based on relevant out-of-sample interval forecasting criteria. The benchmark Taylor rule model is also found to perform relatively better than the standard monetary model and the purchasing power parity (PPP) model.

In a seminal paper, Meese and Rogoff (1983) find that economic fundamentals - such as the money supply, trade balance and national income - are of little use in forecasting exchange rates. They show that existing models cannot forecast exchange rates better than the random walk in terms of out-of-sample forecasting accuracy. This finding suggests that exchange rates may be determined by something purely random rather than economic fundamentals. Meese and Rogoff's (1983) finding has been named the Meese-Rogoff puzzle in the literature.

In defending fundamental-based exchange rate models, various combinations of economic variables and econometric methods have been used in attempts to overturn Meese and Rogoff's finding. For instance, Mark (1995) finds greater exchange rate predictability at longer horizons.[1] Groen (2000) and Mark and Sul (2001) detect exchange rate predictability by using panel data. Kilian and Taylor (2003) find that exchange rates can be predicted from economic models at horizons of 2 to 3 years after taking into account the possibility of nonlinear exchange rate dynamics. Faust, Rogers, and Wright (2003) find that the economic models consistently perform better using real-time data than revised data, though they do not perform better than the random walk.

Recently, there is a growing strand of literature that uses Taylor rules to model exchange rate determination. Engel and West (2005) derive the exchange rate as present-value asset price from a Taylor rule model. They also find a positive correlation between the model-based exchange rate and the actual real exchange rate between the US dollar and the Deutschmark (Engel and West, 2006). Mark (2007) examines the role of Taylor-rule fundamentals for exchange rate determination in a model with learning. In his model, agents use least-square learning rules to acquire information about the numerical values of the model's coefficients. He finds that the model is able to capture six major swings of the real Deutschemark-Dollar exchange rate from 1973 to 2005. Molodtsova and Papell (2008) find significant short-term out-of-sample predictabil-

[1] Chinn and Meese (1995), and MacDonald and Taylor (1994) find similar results. However, the long-horizon exchange rate predictability in Mark (1995) has been challenged by Kilian (1999) and Berkowitz and Giorgianni (2001) in succeeding studies.

ity of exchange rates with Taylor-rule fundamentals for 11 out of 12 currencies vis-á-vis the U.S. dollar over the post-Bretton Woods period. Molodtsova, Nikolsko-Rzhevskyy, and Papell (2008) find evidence of out-of-sample predictability for the dollar/mark nominal exchange rate with forecasts based on Taylor rule fundamentals using real-time data, but not revised data.

Our paper joins the above literature of Taylor-rule exchange rate models. However, we addresses the Meese and Rogoff puzzle from a different perspective: interval forecasting. A forecast interval captures not only the expected future value of the exchange rate, but a range in which the exchange rate may lie with a certain probability, given a set of predictors available at the time of forecast. Our contribution to the literature is twofold. First, we find that for twelve OECD exchange rates, the Taylor rule models in general generate tighter forecast intervals than the random walk, given that their intervals cover the realized exchange rates (statistically) equally well. This finding suggests an intuitive connection between exchange rates and economic fundamentals beyond point forecasting: the use of economic variables as predictors help narrow down the range in which future exchange rates may lie, compared to random walk forecast intervals which are essentially based on unconditional distributions of exchange rates. Second, we propose an inference framework for cross-model comparison of out-of-sample forecast intervals. The proposed framework can be used for forecast-interval evaluation in a broader context, not just for the models and methods used in this paper. For instance, the framework can also be used to evaluate out-of-sample inflation forecasting.

As we will discuss later, we in fact derive forecast intervals from estimates of the distribution of changes in the exchange rate. Hence, in principle, evaluations across models can be done based on distributions instead of forecast intervals. However, focusing on interval forecasting performance allows us to compare models in two dimensions that are more relevant to practitioners: empirical coverage and length.

While the literature on interval forecasting for exchange rates is sparse, several authors have studied out-of-sample exchange rate density forecasts, from which interval forecasts can be derived. Diebold, Hahn and Tay (1999) use the RiskMetrics model of JP Morgan (1996) to compute half-hour-ahead density forecasts for Deutschmark/Dollar and Yen/Dollar returns. Christoffersen and Mazzotta (2005) provide option-implied one-day-ahead density and interval forecasts for four major exchange rates. Boero and Marrocu (2004) obtain one-day-ahead density forecasts for the Euro nominal effective exchange rate using the self-exciting threshold autoregressive (SETAR) models. Sarno and Valente (2005) evaluate the exchange rate density forecasting performance of the Markov-switching vector equilibrium correction model that is developed by Clarida, Sarno, Taylor and Valente (2003). They find that information from the term structure of forward premia help the model to outperform the random walk in forecasting the out-of-sample densities of the spot exchange rate. More recently, Hong, Li and Zhao (2007) construct half-hour-ahead density forecasts for

Euro/Dollar and Yen/Dollar exchange rates using a comprehensive set of univariate time series models that capture fat tails, time-varying volatility and regime switches.

There are several common features across the studies listed above, which make them different from our paper. First, the focus of the above studies is not to make connections between the exchange rate and economic fundamentals. These studies use high frequency data, which are not available for most conventional economic fundamentals. For instance, Diebold, Hahn and Tay (1999) and Hong, Li and Zhao (2007) use intra-day data. With the exception of Sarno and Valente (2005), all the studies focus only on univariate time series models. Second, these studies do not consider multi-horizon-ahead forecasts, perhaps due to the fact that their models are often highly nonlinear. Iterating nonlinear density models multiple horizons ahead is analytically difficult, if not infeasible. Lastly, the above studies assume that the densities are analytically defined for a given model. The semi-parametric method used in this paper does not impose such restrictions.

Our choice of semi-parametric method is motivated by the difficulty of using macroeconomic models in exchange rate interval forecasting: these models typically do not describe the future *distributions* of exchange rates. For instance, the Taylor rule models considered in this paper do not describe any features of the data beyond the conditional means of future exchange rates. We address this difficulty by applying *Robust Semi-parametric* forecast intervals (from hereon *RS* forecast intervals) of Wu (2007).[2] This method is useful since it does not require the model be correctly specified, or contain parametric assumptions about the future distribution of exchange rates.

We apply RS forecast intervals to a set of Taylor rule models that differ in terms of the assumptions on policy and interest rate smoothing rules. Following Molodtsova and Papell (2008), we include twelve OECD exchange rates (relative to the US dollar) over the post-Bretton Woods period in our dataset. For these twelve exchange rates, the out-of-sample RS forecast intervals at different forecast horizons are generated from the Taylor rule models and then compared with those of the random walk. The empirical coverages and lengths of forecast intervals are used as the evaluation criteria. Our empirical coverage and length tests are modified from Giacomini and White's (2006) predictive accuracy tests in the case of rolling but fixed-size estimation samples.

For a given nominal coverage (probability), the empirical coverage of forecast intervals derived from a forecasting model is the probability that the out-of-sample realizations (exchange rates) lie in the intervals. The length of the intervals is a measure of its tightness: the distance between its upper and lower bound. In general, the empirical coverage is not the same as its nominal coverage. Significantly missing the nominal coverage indicates poor quality of the model and intervals. One certainly wants the forecast intervals

[2]For brevity, we omit RS and simply say forecast intervals when we believe that it causes no confusion.

to contain out-of-sample realizations as close as possible to the probability they target. Most evaluation methods in the literature focus on comparing empirical coverages across models, following the seminal work of Christoffersen (1998). Following this literature, we first test whether forecast intervals of the Taylor rule models and the random walk have equally accurate empirical coverages. The model with more accurate coverages is considered the better model. In the cases where equal coverage accuracy cannot be rejected, we further test whether the lengths of forecast intervals are the same. The model with tighter forecast intervals provides more useful information about future values of the data, and hence is considered as a more useful forecasting model.

It is also important to establish what this paper is *not* attempting. First, the inference procedure does not carry the purpose of finding the *correct* model specification. Rather, inference is on how useful models are in generating forecast intervals, measured in terms of empirical coverages and lengths. Second, this paper does not consider the possibility that there might be alternatives to RS forecast intervals for the exchange rate models we consider. Some models might perform better if parametric distribution assumptions (e.g. the forecast errors are conditionally heteroskedastic and $t-$distributed) or other assumptions (e.g. the forecast errors are independent of the predictors) are added. One could presumably estimate the forecast intervals differently based on the same models, and then compare those with the RS forecast intervals, but this is out of the scope of this paper. As we described, we choose RS method for its robustness and flexibility achieved by the semi-parametric approach.

Our benchmark Taylor rule model is from Engel and West (2005) and Engel, Wang, and Wu (2008). For the purpose of comparison, several alternative Taylor rule models are also considered. These setups have been studied by Molodtsova and Papell (2008) and Engel, Mark, and West (2007). In general, we find that the Taylor rule models perform better than the random walk model, especially at long horizons: the models either have more accurate empirical coverages than the random walk, or in cases of equal coverage accuracy, the models have tighter forecast intervals than the random walk. The evidence of exchange rate predictability is much weaker in coverage tests than in length tests. In most cases, the Taylor rule models and the random walk have statistically equally accurate empirical coverages. So, under the conventional coverage test, the random walk model and the Taylor rule models perform equally well. However, the results of length tests suggest that Taylor rule fundamentals are useful in generating tighter forecasts intervals without losing accuracy in empirical coverages.

We also consider two other popular models in the literature: the monetary model and the model of purchasing power parity (PPP). Based on the same criteria, both models are found to perform better than the random walk in interval forecasting. As with the Taylor rule models, most evidence of exchange rate predictability comes from the length test: economic models have tighter forecast intervals than the random walk given statistically equivalent coverage accuracy. The monetary model performs slightly worse than the benchmark Taylor rule model and the PPP model. The benchmark Taylor rule model performs better than the PPP model at short horizons and equally well at long horizons.

Our findings suggest that exchange rate movements are linked to economic fundamentals. However, we acknowledge that the Meese-Rogoff puzzle remains difficult to understand. Although Taylor rule models offer statistically significant length reductions over the random walk, the reduction of length is quantitatively small. Forecasting exchange rates remains a difficult task in practice. There are some impressive advances in the literature, but most empirical findings remain fragile. As mentioned in Cheung, Chinn, and Pascual (2005), forecasts from economic fundamentals may work well for some currencies during certain sample periods but not for other currencies or sample periods. Engel, Mark, and West (2007) recently show that a relatively robust finding is that exchange rates are more predictable at longer horizons, especially when using panel data. We find greater predictability at longer horizons in our exercise. It would be of interest to investigate connections between our findings and theirs.

Several recent studies have attacked the puzzle from a different angle: there are reasons that economic fundamentals cannot forecast the exchange rate, even if the exchange rate is determined by these fundamentals. Engel and West (2005) show that existing exchange rate models can be written in a present-value asset-pricing format. In these models, exchange rates are determined not only by current fundamentals but also by expectations of what the fundamentals will be in the future. When the discount factor is large (close to one), current fundamentals receive very little weight in determining the exchange rate. Not surprisingly, the fundamentals are not very useful in forecasting. Nason and Rogers (2008) generalize the Engel-West theorem to a class of open-economy dynamic stochastic general equilibrium (DSGE) models. Other factors such as parameter instability and mis-specification (for instance, Rossi 2005) may also play important roles in understanding the puzzle. It is interesting to investigate conditions under which we can reconcile our findings with these studies.

The remainder of this paper is organized as follows. Section two describes the forecasting models we use, as well as the data. In section three, we illustrate how the RS forecast intervals are constructed from a given model. We also propose loss criteria to evaluate the quality of the forecast intervals, and test statistics that are based on Giacomini and White (2006). Section four presents results of out-of-sample forecast evaluation. Finally, section five contains concluding remarks.

2 Models and Data

Seven models are considered in this paper. Let $m = 1, 2, ..., 7$ be the index of these models and the first model be the benchmark model. A general setup of the models takes the form of

$$s_{t+h} - s_t = \alpha_{m,h} + \beta'_{m,h}\mathbf{X}_{m,t} + \varepsilon_{m,t+h}, \tag{1}$$

where $s_{t+h} - s_t$ is h-period changes of the (log) exchange rate, and $\mathbf{X}_{m,t}$ contains economic variables that are used in model m. Following the literature of long-horizon regressions, both short- and long-horizon forecasts are considered. Models differ in economic variables that are included in matrix $\mathbf{X}_{m,t}$. In the benchmark model,

$$\mathbf{X}_{1,t} \equiv \left[\begin{array}{ccc} \pi_t - \pi_t^* & y_t^{gap} - y_t^{gap*} & q_t \end{array} \right],$$

where π_t (π_t^*) is the inflation rate, and y_t^{gap} (y_t^{gap*}) is output gap in the home (foreign) country. The real exchange rate q_t is defined as $q_t \equiv s_t + p_t^* - p_t$, where p_t (p_t^*) is the (log) consumer price index in the home (foreign) country. This setup is motivated by the Taylor rule model in Engel and West (2005) and Engel, Wang, and Wu (2008). The next subsection describes this benchmark Taylor rule model in detail.

We also consider the following models that have been studied in the literature:

- Model 2: $\mathbf{X}_{2,t} \equiv \left[\begin{array}{cc} \pi_t - \pi_t^* & y_t^{gap} - y_t^{gap*} \end{array} \right]$
- Model 3: $\mathbf{X}_{3,t} \equiv \left[\begin{array}{ccc} \pi_t - \pi_t^* & y_t^{gap} - y_t^{gap*} & i_{t-1} - i_{t-1}^* \end{array} \right]$, where i_t (i_t^*) is the short-term interest rate in the home (foreign) country.
- Model 4: $\mathbf{X}_{4,t} \equiv \left[\begin{array}{cccc} \pi_t - \pi_t^* & y_t^{gap} - y_t^{gap*} & q_t & i_{t-1} - i_{t-1}^* \end{array} \right]$
- Model 5: $\mathbf{X}_{5,t} \equiv q_t$
- Model 6: $\mathbf{X}_{6,t} \equiv \left[\begin{array}{c} s_t - [(m_t - m_t^*) - (y_t - y_t^*)] \end{array} \right]$, where m_t (m_t^*) is the money supply and y_t (y_t^*) is total output in the home (foreign) country.
- Model 7: $\mathbf{X}_{7,t} \equiv 0$

Models 2-4 are the Taylor rule models studied in Molodtsova and Papell (2008). Model 2 can be considered as the constrained benchmark model in which PPP always holds. Molodtsova and Papell (2008) include interest rate lags in models 3 and 4 to take into account potential interest rate smoothing rules of the central bank. Model 5 is the purchasing power parity (PPP) model and model 6 is the monetary model. Both models have been widely used in the literature. See Molodtsova and Papell (2008) for the PPP model and Mark (1995) for the monetary model. Model 7 is the driftless random walk model ($\alpha_{7,h} \equiv 0$).[3] Given a date τ and horizon h, the objective is to estimate the forecast distribution of $s_{\tau+h} - s_\tau$ conditional on $\mathbf{X}_{m,\tau}$, and subsequently build forecast intervals from the estimated forecast distribution. Before moving to the econometric method, we first describe the Taylor rule model that motivates the setup of our benchmark model.

2.1 Benchmark Taylor Rule Model

Our benchmark model is the Taylor rule model that is derived in Engel and West (2005) and Engel, Wang, and Wu (2008). Following Molodtsova and Papell (2008), we focus on models that depend only on *current* levels of inflation and output gap.[4] The Taylor rule in the home country takes the form of

$$\bar{i}_t = \bar{i} + \delta_\pi(\pi_t - \bar{\pi}) + \delta_y y_t^{gap} + u_t, \tag{2}$$

where $\bar{i}_t$ is the central bank's target for short-term interest rate at time t, $\bar{i}$ is the equilibrium long-run rate, π_t is the inflation rate, $\bar{\pi}$ is the target inflation rate, and y_t^{gap} is output gap. The foreign country is assumed to follow a symmetric Taylor rule. In addition, we follow Engel and West (2005) to assume that the foreign country targets the exchange rate in its Taylor rule:

$$\bar{i}_t^* = \bar{i} + \delta_\pi(\pi_t^* - \bar{\pi}) + \delta_y y_t^{gap*} + \delta_s(s_t - \bar{s}_t) + u_t^*, \tag{3}$$

where $\bar{s}_t$ is the targeted exchange rate. Assume that the foreign country targets the PPP level of the exchange rate: $\bar{s}_t = p_t - p_t^*$, where p_t and p_t^* are logarithms of the home and foreign aggregate prices. In equation (3), we assume that the policy parameters take the same values in the home and foreign countries. To simplify our presentation, we assume that the home and foreign countries have the same long-run inflation and interest rates. Such restrictions have been relaxed in our econometric model after we include a constant term in estimations.

We do not consider interest rate smoothing in our benchmark model. That is, the actual interest rate

[3] We also tried the random walk with a drift. It does not change our results.

[4] Clarida, Gali, and Gertler (1998) find empirical support for forward looking Taylor rules. *Forward looking* Taylor rules are ruled out because they require forecasts of predictors, which creates additional complications in out-of-sample forecasting.

(i_t) is identical to the target rate in the benchmark model:

$$i_t = \bar{i}_t. \tag{4}$$

Molodtsova and Papell (2008) consider the following interest rate smoothing rule:

$$i_t = (1-\rho)\bar{i}_t + \rho i_{t-1} + \nu_t, \tag{5}$$

where ρ is the interest rate smoothing parameter. We include these setups in models 3 and 4. Note that our estimation methods do not require the monetary policy shock u_t and the interest rate smoothing shock ν_t to satisfy any assumptions, aside from smoothness of their distributions when conditioned on predictors.

Substituting the difference of equations (2) and (3) to Uncovered Interest-rate Parity (UIP), we have

$$s_t = E_t \left\{ (1-b)\sum_{j=0}^{\infty} b^j (p_{t+j} - p^*_{t+j}) - b\sum_{j=0}^{\infty} b^j \left[\delta_y (y^{gap}_{t+j} - y^{gap*}_{t+j}) + \delta_\pi (\pi_{t+j} - \pi^*_{t+j})\right] \right\}, \tag{6}$$

where the discount factor $b = \frac{1}{1+\delta_s}$. Under some conditions, the present value asset pricing format in equation (6) can be written into an error-correction form:[5]

$$s_{t+h} - s_t = \alpha_h + \beta_h z_t + \varepsilon_{t+h}, \tag{7}$$

where the deviation of the exchange rate from its equilibrium level is defined as:

$$z_t = s_t - p_t + p^*_t + \frac{b}{1-b}\left[\delta_y (y^{gap}_t - y^{gap*}_t) + \delta_\pi (\pi_t - \pi^*_t)\right]. \tag{8}$$

We use equation (7) as our benchmark setup in calculating h-horizon-ahead out-of-sample forecasting intervals. According to equation (8), the matrix $\mathbf{X}_{1,t}$ in equation (1) includes economic variables $q_t \equiv s_t + p^*_t - p_t$, $y^{gap}_t - y^{gap*}_t$, and $\pi_t - \pi^*_t$.

[5]See appendix for more detail. While the long-horizon regression format of the benchmark Taylor model is derived directly from the underlying Taylor rule model, this is not the case for the models with interest rate smoothing (models 3 and 4). Molodtsova and Papell (2007) only consider the short-horizon regression for the Taylor rule models. We include long-horizon regressions of these models only for the purpose of comparison.

2.2 Data

The forecasting models and the corresponding forecast intervals are estimated using monthly data for twelve OECD countries. The United States is treated as the foreign country in all cases. For each country we synchronize the beginning and end dates of the data across all models estimated. The twelve countries and periods considered are: Australia (73:03-06:6), Canada (75:01-06:6), Denmark (73:03-06:6), France (77:12-98:12), Germany (73:03-98:12), Italy (74:12-98:12), Japan (73:03-06:6), Netherlands (73:03-98:12), Portugal (83:01-98:12), Sweden (73:03-04:11), Switzerland (75:09-06:6), and the United Kingdom (73:03-06:4).

The data is taken from Molodtsova and Papell (2008).[6] With the exception of interest rates, the data is transformed by taking natural logs and then multiplied by 100. The nominal exchange rates are end-of-month rates taken from the Federal Reserve Bank of St. Louis database. Output data y_t are proxied by Industrial Production (IP) from the International Financial Statistics (IFS) database. IP data for Australia and Switzerland are only available at quarterly frequency, and hence are transformed from quarterly to monthly observations using the quadratic-match average option in Eviews 4.0 by Molodtsova and Papell (2008). Following Engel and West (2006), the output gap y_t^{gap} is calculated by quadratically de-trending the industrial production for each country.

Prices data p_t are proxied by Consumer Price Index (CPI) from the IFS database. Again, CPI for Australia is only available at quarterly frequency and quadratic-match average is used to impute monthly observations. Inflation rates are calculated by taking the first differences of the logs of CPIs. Money market rate from IFS (or "call money rate") is used as a measure of the short-term interest rate set by the central bank. Finally, M1 is used to measure the money supply for most countries. M0 for the UK, and M2 for Italy and Netherlands is used due to the unavailability of M1 data.

3 Econometric Method

For a given model m, the objective is to estimate from equation (1) the distribution of $s_{\tau+h} - s_\tau$ conditional on data $\mathbf{X}_{m,\tau}$ that is observed up to time τ. This is the h-horizon-ahead *forecast distribution* of the exchange rate, from which the corresponding *forecast interval* can be derived. For a given α, the forecast interval of coverage $\alpha \in (0, 1)$ is an interval in which $s_{\tau+h} - s_\tau$ is supposed to lie with a probability of α.

Models $m = 1, ..., 7$ in equation (1) provide only point forecasts of $s_{\tau+h} - s_\tau$. In order to construct forecast intervals for a given model, we apply robust semi-parametric (RS) forecast intervals to all models.

[6] We thank the authors for the data, which we downloaded from David Papell's website. For the exact line numbers and sources of the data, see the data appendix of Molodtsova and Papell (2008).

The nominal α-coverage forecast interval of $s_{\tau+h} - s_\tau$ conditional on $\mathbf{X}_{m,\tau}$ can be obtained by the following three-step procedure:

Step 1. Estimate model m by OLS and obtain residuals $\widehat{\varepsilon}_{m,t+h} \equiv s_{t+h} - s_t - \widehat{\alpha}_{m,h} + \widehat{\beta}'_{m,h}\mathbf{X}_{m,t}$, for $t = 1, ..., \tau - h$.

Step 2. For a range of values of ε (sorted residuals $\{\widehat{\varepsilon}_{m,t+h}\}_{t=1}^{\tau-h}$), estimate the conditional distribution of $\varepsilon_{m,\tau+h}|\mathbf{X}_{m,\tau}$ by

$$\widehat{P}(\varepsilon_{m,\tau+h} \leq \varepsilon|\mathbf{X}_{m,\tau}) \equiv \frac{\sum_{t=1}^{\tau-h} 1(\widehat{\varepsilon}_{m,t+h} \leq \varepsilon)\mathbf{K}_b(\mathbf{X}_{m,t} - \mathbf{X}_{m,\tau})}{\sum_{t=1}^{\tau-h} \mathbf{K}_b(\mathbf{X}_{m,t} - \mathbf{X}_{m,\tau})}, \tag{9}$$

where $\mathbf{K}_b(\mathbf{X}_{m,t} - \mathbf{X}_{m,\tau}) \equiv b^{-d}\mathbf{K}((\mathbf{X}_{m,t} - \mathbf{X}_{m,\tau})/b)$, $\mathbf{K}(\cdot)$ is a multivariate Gaussian kernel with dimension same as that of $\mathbf{X}_{m,t}$, and b is the smoothing parameter or bandwidth.[7]

Step 3. Find the $(1-\alpha)/2$ and $(1+\alpha)/2$ quantiles of the estimated distribution, which are denoted by $\widehat{\varepsilon}_{m,h}^{(1-\alpha)/2}$ and $\widehat{\varepsilon}_{m,h}^{(1+\alpha)/2}$. The estimate of the α-coverage forecast interval for $s_{\tau+h} - s_\tau$ conditional on $\mathbf{X}_{m,\tau}$ is

$$\widehat{I}_{m,\tau+h}^{\alpha} \equiv (\widehat{\beta}'_{m,h}\mathbf{X}_{m,\tau} + \widehat{\varepsilon}_{m,h}^{(1-\alpha)/2}, \widehat{\beta}'_{m,h}\mathbf{X}_{m,\tau} + \widehat{\varepsilon}_{m,h}^{(1+\alpha)/2}) \tag{10}$$

For each model m, the above method uses the forecast models in equation (1) to estimate the location of the forecast distribution, while nonparametric kernel distribution estimate is used to estimate the shape. As a result, the interval obtained from this method is *semi-parametric*. Wu (2007) shows that under some weak regularity conditions, this method always consistently estimates the forecast distribution,[8] and hence the forecast intervals, of $s_{\tau+h} - s_\tau$ conditional on $\mathbf{X}_{m,\tau}$, regardless of the quality of model m. That is, the forecast intervals are *robust*. Stationarity of economic variables is one of those regularity conditions. In our models, exchange rate differences, interest rates and inflation rates are well-known to be stationary, while empirical tests for real exchange rates and output gaps generate mixed results. These results may be driven by the difficulty of distinguishing a stationary but persistent variable from a non-stationary one. In this paper, we take the stationarity of these variables as given.

Model seven is the random walk model. The estimator in equation (9) becomes the Empirical Distribution Function (EDF) of the exchange rate innovations. Under regularity conditions, equation (9) consistently estimates the unconditional distribution of $s_{\tau+h} - s_\tau$, and can be used to form forecast intervals for $s_{\tau+h}$.

[7]We choose b using the method of Hall, Wolff, and Yao (1999).

[8]It is consistent in the sense of convergence in probability as the estimation sample size goes to infinity.

The forecast intervals of economic models and the random walk are compared. Our interest is to test whether RS forecast intervals based on economic models are more accurate than those based on the random walk model. We focus on the empirical coverage and the length of forecast intervals in our tests.

Following Christoffersen (1998) and related work, the first standard we use is the empirical coverage. The empirical coverage should be as close as possible to the nominal coverage (α). Significantly missing the nominal coverage indicates the inadequacy of the model and predictors for the given sample size. For instance, if 90% forecast intervals calculated from a model contain only 50% of out-of-sample observations, the model can hardly be identified as useful for interval forecasting. This case is called under-coverage. In contrast, over-coverage implies that the intervals could be reduced in length (or improved in tightness), but the forecast interval method and model are unable to do that for the given sample size. An economic model is said to outperform the random walk if its empirical coverage is more accurate than that of the random walk.

On the other hand, the empirical coverage of an economic model may be equally accurate as that of the random walk model, but the economic model has tighter forecast intervals than the random walk. We argue that the lengths of forecast intervals signify the informativeness of the intervals given that these intervals have equally accurate empirical coverages. In this case, the economic model is also considered to outperform the random walk in forecasting exchange rates. The empirical coverage and length tests are conducted at both short and long horizons for six economic models relative to the random walk for each of the twelve OECD exchange rates.

We use tests that are applications of the unconditional predictive accuracy inference framework of Giacomini and White (2006). Unlike the tests of Diebold and Mariano (1995) and West (1996), our forecast evaluation tests do not focus the asymptotic features of the forecasts. Rather, in the spirit of Giacomini and White (2006), we are comparing the population features of forecasts generated by rolling samples of fixed sample size. This contrasts the traditional forecast evaluation methods in that although it uses asymptotic approximations to do the testing, the inference is not on the asymptotic properties of forecasts, but on their population *finite sample properties*. We acknowledge that the philosophy of this inference framework remains a point of contention, but it does tackle three important evaluation difficulties in this paper. First, it allows for evaluation of forecast intervals that are not parametrically derived. The density evaluation methods developed in well-known studies such as Diebold, Gunther, Tay (1998), Corradi and Swanson (2006a) and references within Corradi and Swanson (2006b) require that the forecast distributions be parametrically specified. Giacomini and White's (2006) method overcomes this challenge by allowing comparisons among parametric, semi-parametric and nonparametric forecasts. As a result, in the cases of semi-parametric and nonparametric forecasts, it also allows comparison of models with predictors of different dimensions, as evident in our exercise. Second, by comparing finite sample properties of RS forecast intervals derived from different models, we avoid rejecting models that are mis-specified,[9] but are nonetheless good approximations useful for forecasting. Finally, we can individually (though not jointly) test whether the forecast intervals differ in terms of empirical coverages and lengths, for the given estimation sample, and not confined to focus only on empirical coverages or holistic properties of forecast distribution such as probability integral transform.

3.1 Test of Equal Empirical Coverages

Suppose the sample size available to the researcher is T and all data are collected in a vector $\mathbf{W}_t$. Our inference procedure is based on a rolling estimation scheme, with the size of the rolling window fixed while $T \to \infty$. Let $T = R + N$ and R be the size of the rolling window. For each horizon h and model m, a sequence of $N(h) = N + 1 - h$ α-coverage forecast intervals are generated using rolling data: $\{\mathbf{W}_t\}_{t=1}^{R}$ for forecast for date $R + h$, $\{\mathbf{W}_t\}_{t=2}^{R+1}$ for forecast for date $R + h + 1$, and so on, until forecast for date T is generated using $\{\mathbf{W}_t\}_{t=N(h)}^{R+N(h)-1}$.

Under this fixed-sample-size rolling scheme, for each finite h we have $N(h)$ observations to compare the empirical coverages and lengths across m models ($m = 1, 2, ..., 7$). By fixing R, we allow the finite sample properties of the forecast intervals to be preserved as $T \to \infty$. Thus, the forecast intervals and the associated forecast losses are simply functions of a finite and fixed number of random variables. We are interested in approximating the population moments of these objects by taking $N(h) \to \infty$. A loose analogy would be finding the finite-sample properties of a certain parameter estimator when sample size is fixed at R, by a bootstrap with an arbitrarily large number of bootstrap replications.

We conduct individual tests for the empirical coverages and lengths. In each test, we define a corresponding forecast loss, propose a test statistic and derive its asymptotic distribution. As defined in equation (10), let $\widehat{I}^{\alpha}_{m,\tau+h}$ be the $h-$horizon ahead RS forecast interval of model m with a nominal coverage of α. For out-of-sample forecast evaluation, we require $\widehat{I}^{\alpha}_{m,\tau+h}$ to be constructed using data from $t = \tau - R + 1$ to $t = \tau$. The *coverage accuracy loss* is defined as

$$CL^{\alpha}_{m,h} = \left[P(Y_{\tau+h} \in \widehat{I}^{\alpha}_{m,\tau+h}) - \alpha\right]^2. \tag{11}$$

For economic models ($m = 1, ..., 6$), the goal is to compare the coverage accuracy loss of RS forecast intervals

[9]While RS intervals remedy mis-specifications asymptotically, it does not guarantee such corrections in a given finite sample.

of model m with that of the random walk ($m = 7$). The null and alternative hypotheses are:

$$H_0 \quad : \quad \Delta CL^{\alpha}_{m,h} \equiv CL^{\alpha}_{7,h} - CL^{\alpha}_{m,h} = 0$$
$$H_A \quad : \quad \Delta CL^{\alpha}_{m,h} \neq 0.$$

Define the sample analog of the coverage accuracy loss in equation (11):

$$\widehat{CL}^{\alpha}_{m,h} = \left(N(h)^{-1} \sum_{\tau=R}^{T-h} 1(Y_{\tau+h} \in \widehat{I}^{\alpha}_{m,\tau+h}) - \alpha \right)^2,$$

where $1(Y_{\tau+h} \in \widehat{I}^{\alpha}_{m,\tau+h})$ is an index function that equals one when $Y_{\tau+h} \in \widehat{I}^{\alpha}_{m,\tau+h}$, and equals zero otherwise. Applying the asymptotic test of Giacomini and White (2006) to the sequence $\{1(Y_{\tau+h} \in \widehat{I}^{\alpha}_{m,\tau+h})\}_{\tau=R}^{T-h}$ and applying the Delta method, we can show that

$$\sqrt{N(h)}(\Delta\widehat{CL}^{\alpha}_{m,h} - \Delta CL^{\alpha}_{m,h}) \xrightarrow{d} N(0, \Gamma^{'}_{m,h}\Omega_{m,h}\Gamma_{m,h}), \tag{12}$$

where $\xrightarrow{d}$ denotes convergence in distribution, and $\Omega_{m,h}$ is the long-run covariance matrix between $1(Y_{\tau+h} \in \widehat{I}^{\alpha}_{m,\tau+h})$ and $1(Y_{\tau+h} \in \widehat{I}^{\alpha}_{7,\tau+h})$. The matrix $\Gamma_{m,h}$ is defined as:

$$\Gamma_{m,h} \equiv \left[\begin{array}{cc} 2\left(P\left(Y_{\tau+h} \in \widehat{I}^{\alpha}_{m,\tau+h}\right) - \alpha\right) & 2\left(P\left(Y_{\tau+h} \in \widehat{I}^{\alpha}_{7,\tau+h}\right) - \alpha\right) \end{array} \right]'.$$

$\Gamma_{m,h}$ can be estimated consistently by its sample analog $\widehat{\Gamma}_{m,h}$, while $\Omega_{m,h}$ can be estimated by some HAC estimator $\widehat{\Omega}_{m,h}$, such as Newey and West (1987).[10] The test statistic for coverage test is defined as:

$$Ct^{\alpha}_{m,h} \equiv \frac{\sqrt{N(h)}\Delta\widehat{CL}^{\alpha}_{m,h}}{\sqrt{\widehat{\Gamma}^{'}_{m,h}\widehat{\Omega}_{m,h}\widehat{\Gamma}_{m,h}}} \xrightarrow{d} N(0,1) \tag{13}$$

3.2 Test of Equal Empirical Lengths

Define the *length loss* as:

$$LL^{\alpha}_{m,h} \equiv E\left[leb\left(\widehat{I}^{\alpha}_{m,\tau+h}\right)\right], \tag{14}$$

[10]We use Newey and West (1987) for our empirical work, with a window width of 12.

where $leb(\cdot)$ is the Lesbesgue measure. To compare the length loss of RS forecast intervals of economic models $m = 1, 2, ..., 6$ with that of the random walk ($m = 7$), the null and alternative hypotheses are:

$$\begin{aligned} H_0 &: \Delta LL^{\alpha}_{m,h} \equiv LL^{\alpha}_{7,h} - LL^{\alpha}_{m,h} = 0 \\ H_A &: \Delta LL^{\alpha}_{m,h} \neq 0. \end{aligned}$$

The sample analog of the length loss for model m is defined as:

$$\widehat{LL}^{\alpha}_{m,h} = N(h)^{-1} \sum_{\tau=R}^{T-h} leb(\widehat{I}^{\alpha}_{m,\tau+h}).$$

Directly applying the test of Giacomini and White (2006), we have

$$\sqrt{N(h)}(\Delta\widehat{LL}^{\alpha}_{m,h} - \Delta LL^{\alpha}_{m,h}) \xrightarrow{d} N(0, \Sigma_{m,h}), \tag{15}$$

where $\Sigma_{m,h}$ is the long-run variance of $leb\left(\widehat{I}^{\alpha}_{7,\tau+h}\right) - leb\left(\widehat{I}^{\alpha}_{m,\tau+h}\right)$. Let $\widehat{\Sigma}_{m,h}$ be the HAC estimator of $\Sigma_{m,h}$. The test statistic for empirical length is defined as:

$$Lt^{\alpha}_{m,h} \equiv \frac{\sqrt{N(h)}\Delta\widehat{LL}^{\alpha}_{m,h}}{\sqrt{\widehat{\Sigma}_{m,h}}} \xrightarrow{d} N(0, 1). \tag{16}$$

3.3 Discussion

The coverage accuracy loss function is symmetric in our paper. In practice, an asymmetric loss function may be better when looking for an exchange rate forecast model to help make policy or business decisions. Under-coverage is arguably a more severe problem than over-coverage in practical situations. However, the focus of this paper is the disconnect between economic fundamentals and the exchange rate. Our goal is to investigate which model comes closer to the data: the random walk or fundamental-based models. It is not critical in this case whether coverage inaccuracy comes from the under- or over-coverage. We acknowledge that the use of symmetric coverage loss remains a caveat, especially since we are using the coverage accuracy test as a pre-test for the tests of length. Clearly, there is a tradeoff between the empirical coverage and the length of forecast intervals. Given the same center,[11] intervals with under-coverage have shorter lengths than intervals with over-coverage. In this case, the length test is in favor of models that systematically under-cover the targeted nominal coverage when compared to a model that systematically over-covers. This problem

[11] Center here means the half way point between the upper and lower bound for a given interval.

cannot be detected by the coverage accuracy test with symmetric loss function because over- and under-coverage are treated equally. However, our results in section 4 show that there is no evidence of systematic under-coverage for the economic models considered in this paper. For instance, in Table 1, one-month-ahead ($h = 1$) forecast interval over-covers the nominal coverage (90%) for nine out of twelve exchange rates.[12] Note that under-coverage does not guarantee shorter intervals either in our paper, because forecast intervals of different models usually have different centers.[13]

As we have mentioned, comparisons across models can also be done at the distribution level. We choose interval forecasts for two reasons. First, interval forecasts have been widely used and reported by the practitioners. For instance, the Bank of England calculates forecast intervals of inflation in its inflation reports. Second, compared to evaluation metrics for density forecasts, the empirical coverage and length loss functions of interval forecasts, and the subsequent interpretations of test rejection/acceptance are more intuitive.

4 Results

We apply RS forecast intervals for each model for a given nominal coverage of $\alpha = 0.9$. There is no particular reason why we chose 0.9 as the nominal coverage. Some auxiliary results show that our qualitative findings do not depend on the choice of α. Due to different sample sizes across countries, we choose different sizes for the rolling window (R) for different countries. Our rule is very simple: for countries with $T \geq 300$, we choose $R = 200$, otherwise we set $R = 150$.[14] Again, from our experience, tampering with R does not change the qualitative results unless R is chosen to be unusually big or small.

For time horizons $h = 1, 3, 6, 9, 12$ and models $m = 1, ..., 7$, we construct a sequence of $N(h)$ 90% forecast intervals $\{\widehat{I}^{0.9}_{m,\tau+h}\}^{T-h}_{\tau=R}$ for the h-horizon change of the exchange rate $s_{t+h} - s_t$. Then we compare economic models and the random walk by computing empirical coverages, lengths and test statistics $Ct^{0.9}_{m,h}$ and $Lt^{0.9}_{m,h}$ as described in section 3. We first report the results of our benchmark model. After that, results of alternative models are reported and discussed.

[12] These nine exchange rates are the Danish Kroner, the French Franc, the Deutschmark, the Italian Lira, the Japanese Yen, the Dutch Guilder, the Portuguese Escudo, the Swiss Franc, and the British pound. Similar results hold at other horizons.

[13] When comparing the intervals for $S_{\tau+h} - S_\tau$, the random walk model builds the forecast interval around 0, while economic model m build it around $\widehat{\beta}'_{m,h}\mathbf{X}_{m,\tau}$.

[14] The only exception is Portugal, where only 192 data points were available. In this case, we choose R = 120.

4.1 Results of Benchmark Model

Table 1 shows results of the benchmark Taylor rule model. For each time horizon h and exchange rate, the first column (Cov.) reports the empirical coverage for the given nominal coverage of 90%. The second column (Leng.) reports the length of forecast intervals (the distance between upper and lower bounds). The length is multiplied by 100 and therefore expressed in terms of the percentage change of the exchange rate. For instance, the length of the one-month-ahead forecast interval for the Australian dollar is 7.114. On average, the distance between the upper and lower bound of the one-month-ahead forecast interval for the Australian dollar is 7.114% change of the Australian dollar against the US dollar. We use superscripts a, b, and c to denote that the null hypothesis of equal empirical coverage accuracy/length is rejected in favor of the Taylor rule model at a confidence level of 10%, 5%, and 1% respectively. Superscripts x, y, and z are used for rejections in favor of the random walk analogously.

We summarize our findings in three panels. In the first panel (*(1) Coverage Test*), the row of "Model Better" reports the number of exchange rates that the Taylor rule model has more accurate empirical coverages than the random walk. The row of "RW Better" reports the number of exchange rates for which the random walk outperforms the Taylor rule model under the same criterion. In the second panel (*(2) Length Test Given Equal Coverage Accuracy*), a better model is the one with tighter forecast intervals given equal coverage accuracy. In the last panel (*(1)+(2)*), a better model is the one with either more accurate coverages, or tighter forecast intervals given equal coverage accuracy.

For most exchange rates and time horizons, the Taylor rule model and the random walk model have statistically equally accurate empirical coverages. The null hypothesis of equal coverage accuracy is rejected in only nine out of sixty tests (one rejection at horizon 3, two rejections at horizons 6 and 9, and four at horizon 12). All nine rejections are in favor of the Taylor rule model. That is, the empirical coverage of the Taylor rule model is closer to the nominal coverage than those of the random walk. However, based on the number of rejections (9) in a total of sixty tests, there is no strong evidence that the Taylor rule model can generate more accurate empirical coverages than the random walk. The coverage tests at horizon twelve have more rejections in favor of the Taylor rule model than that at short horizons ($h = 1, 3$). However, this pattern in coverage tests does not exist in other models that will be discussed in next subsection.

In cases where the Taylor rule model and the random walk have equally accurate empirical coverages, the Taylor rule model generally has equal or significantly tighter forecast intervals than the random walk. In forty-one out of fifty-one cases, the null hypothesis of equally tight forecast intervals is rejected in favor of the Taylor rule model. In contrast, the null hypothesis is rejected in only two cases in favor of the random walk.

The evidence of exchange rate predictability is more pronounced at longer horizons. At horizons nine and twelve ($h = 9, 12$), for cases where empirical coverage accuracies between the random walk and the Taylor rule model are statistically equivalent, the Taylor rule model has significantly tighter forecast intervals than the random walk.

As for each individual exchange rate, the benchmark Taylor rule model works best for the Canadian dollar, the French Franc, the Deutschmark, and the Swedish Krona: for all time horizons, the model has tighter forecast intervals than the random walk, while their empirical coverages are statistically equally accurate. The Taylor rule model performs better than the random walk in most horizons for remaining exchange rates except the Japanese yen, for which the Taylor rule model outperforms the random walk only at long horizons.

4.2 Results of Alternative Models

Five alternative economic models are also compared with the random walk: three alternative Taylor rule models that are studied in Molodtsova and Papell (2008), the PPP model, and the monetary model. Tables 2-6 report results of these alternative models.

In general, results of coverage tests do not show strong evidence that economic models can generate more accurate coverages than the random walk at either short or long horizons. Though the benchmark Taylor rule model shows a sign of long-horizon predictability based on coverage accuracy tests, there is no clear evidence for such a pattern in any other models. However, after considering length tests, we find that economic models perform better than the random walk, especially at long horizons. The Taylor rule model four (Table 4) and the PPP model (Table 5) perform the best among alternative models. Results of these two models are very similar to that of the benchmark Taylor rule model. At horizon twelve, both models outperform the random walk for all twelve exchange rates under our out-of-sample forecast interval evaluation criteria. The performance of the Taylor rule model two (Table 2) and three (Table 3) is relatively less impressive than other models, but still for about half of exchange rates, the economic models outperform the random walk at several horizons in out-of-sample interval forecasts.

Comparing the benchmark Taylor rule model, the PPP model and the monetary model, the performance of the monetary model (Table 6) is slightly worse than the other two models at long horizons. Compared to the Taylor rule and PPP models, the monetary model outperforms the random walk for a smaller number of exchange rates at horizons 6, 9, and 12. The Taylor rule model and the monetary model perform relatively better than the PPP model at short horizons. Overall, the benchmark Taylor rule model seems to perform slightly better than the monetary and PPP models. Molodtsova and Papell (2008) find similar results in their point forecasts.

4.3 Discussion

After Mark (1995) first documents exchange rate predictability at long horizons, long-horizon exchange rate predictability has become a very active area in the literature. With panel data, Engel, Mark, and West (2007) recently show that the long-horizon predictability of the exchange rate is relatively robust in the exchange rate forecasting literature. We find similar results in our interval forecasts. The evidence of long-horizon predictability seems robust across different models and currencies when both empirical coverage and length tests are used. At horizon twelve, all economic models outperform the random walk for seven exchange rates: the Australia dollar, Canadian dollar, Italian Lira, Japanese yen, Portuguese escudo, Swedish krona, and the British pound in the sense that interval lengths of economic models are smaller than those of the random walk, given equivalent coverage accuracy. This is true only for the Italian Lira at horizon one. We also notice that there is no clear evidence of long-horizon predictability based on the tests of empirical coverage accuracy only.

Molodtsova and Papell (2008) find strong out-of-sample exchange rate predictability for Taylor rule models even at the short horizon. In our paper, the evidence for exchange rate predictability at short horizons is not very strong. This finding may be a result of some assumptions we have used to simplify our computation. For instance, an α-coverage forecast interval will always be constructed using the $(1-\alpha)/2$ and $(1+\alpha)/2$ quantiles. Alternatively, we can choose quantiles that minimize the length of intervals, given the nominal coverage.[15] We have also assumed symmetric Taylor rules. Relaxing these assumptions may help us find exchange rate predictability at short horizons. In addition, the development of more powerful testing methods may also be helpful. The evidence of exchange rate predictability in Molodtsova and Papell (2008) is partly driven by the testing method recently developed by Clark and West (2006, 2007). We acknowledge that whether or not short-horizon results can be improved remains an interesting question, but do not pursue this in the current paper. The purpose of this paper is to show the connection between the exchange rate and economic fundamentals from an interval forecasting perspective. Predictability either at short- or long-horizons will serve this purpose.

Though we find that economic fundamentals are helpful for forecasting exchange rates, we acknowledge that exchange rate forecasting in practice is still a difficult task. The forecast intervals from economic models are statistically tighter than those of the random walk, but they remain fairly wide. For instance, the distance

[15] See Wu (2007) for more discussions.

between the upper and lower bound of three-month-ahead forecast intervals is usually a 20% change of the exchange rates. Figures 1-3 show forecast intervals generated by the benchmark Taylor rule model and the random walk for the British pound, the Deutschmark, and the Japanese yen at different horizons.[16] To facilitate graphical comparisons, the 6- and 12-month-ahead forecast intervals of the random walk have been re-centered so that they have the same center as the forecast intervals of the Taylor rule model. In these figures, the Taylor rule model has tighter forecast intervals, especially at the horizon of 12 months, than the random walk. However, the difference is quantitatively small.

5 Conclusion

There is a growing strand of literature that uses Taylor rules to model exchange rate movements. Our paper contributes to the literature by showing that Taylor rule fundamentals are useful in forecasting distribution of exchange rates. We apply Robust Semi-parametric forecast intervals of Wu (2007) to a group of Taylor models for twelve OECD exchange rates. The forecast intervals generated by the Taylor rule models are in general tighter than those of the random walk, given that these intervals cover the realized exchange rates equally well. The evidence of exchange rate predictability is more pronounced at longer horizons, a result that echoes previous long-horizon studies such as Mark (1995). The benchmark Taylor rule model is also found to perform better than the monetary and PPP models based on out-of-sample interval forecasts.

Though we find some empirical support for the connection between the exchange rate and economic fundamentals, we acknowledge that the detected connection is weak. The reductions of the lengths of forecast intervals are quantitatively small, though they are statistically significant. Forecasting exchange rates remains a difficult task in practice. Engel and West (2005) argue that as the discount factor gets closer to one, present value asset pricing models place greater weight on future fundamentals. Consequently, current fundamentals have very weak forecasting power and exchange rates appear to follow approximately a random walk. Under standard assumptions in Engel and West (2005), the Engel-West theorem does not imply that exchange rates are more predictable at longer horizons, or economic models can outperform the random walk in forecasting exchange rates based on out-of-sample interval forecasts. However, modifications to these assumptions may be able to reconcile the Engel-West explanation with empirical findings of exchange rate predictability. For instance, Engel, Wang, and Wu (2008) find that when there exist stationary but persistent unobservable fundamentals, for example risk premium, the Engel-West explanation predicts long-horizon exchange rate predictability in *point forecasts*, though the exchange rate still approximately follows

[16]Figures in other countries show similar patterns. Results are available upon request.

a random walk at short horizons. It would also be of interest to study conditions under which our findings in *interval forecasts* can be reconciled with the Engel-West theorem.

We believe other issues, such as parameter instability (Rossi, 2005), nonlinearity (Kilian and Taylor, 2003), real time data (Faust, Rogers, and Wright, 2003, Molodtsova, Nikolsko-Rzhevskyy, and Papell, 2008), are all contributing to the Meese-Rogoff puzzle. Panel data are also found helpful in detecting exchange rate predictability, especially at long horizons. For instance, see Mark and Sul (2001) and Engel, Mark, and West (2007). It would be interesting to incorporate these studies into interval forecasting. We leave these extensions for future research.

Table 1: Results of Benchmark Taylor Rule Model

	$h=1$		$h=3$		$h=6$		$h=9$		$h=12$	
	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.
Australian Dollar	0.895	7.114	0.888	14.283^c	0.923	20.286^c	0.927	24.468^c	0.915	28.016^c
Canadian Dollar	0.808	3.446^c	0.789	6.351^c	0.808	8.586^c	0.817	10.321^c	0.801	12.614^c
Danish Kroner	0.920	8.668^c	0.939	17.415^c	0.949	26.087	0.969	30.776^c	0.968	36.962^c
French Franc	0.912	8.921^c	0.920	17.674^c	0.928^c	26.007^c	0.957	29.924^c	0.934	36.883^c
Deutschmark	0.927	8.851^c	0.879	18.634^c	0.894	27.923^c	0.960^a	33.734^c	0.969	38.374^c
Italian Lira	0.906	8.754^c	0.875	18.305	0.910	26.788^c	0.862	34.785^c	0.890^a	39.958^c
Japanese Yen	0.915	9.633^z	0.909	19.765	0.902	28.497^c	0.932	33.793^c	0.883	37.333^c
Dutch Guilder	0.917	8.821	0.907	18.649^c	0.933	27.649^c	0.951^a	31.117^c	0.959^a	40.737^c
Portuguese Escudo	0.901	8.205^z	0.913^a	17.899	0.879^c	22.431^c	0.825	25.959^c	0.883^c	32.464^c
Swedish Krona	0.844	7.448^c	0.860	15.405^c	0.874	23.930^c	0.861	30.827^c	0.834	37.432^c
Swiss Franc	0.935	9.759^c	0.946	20.036	0.982	27.682^c	0.994	32.837^c	0.956	38.728^c
British Pound	0.919	8.429	0.923	16.570^c	0.906	23.623^c	0.884	27.849^c	0.903^c	30.814^c
(1) Coverage Test†										
Model Better	0		1		2		2		4	
RW Better	0		0		0		0		0	
(2) Length Test Given Equal Coverage Accuracy‡										
Model Better	7		8		9		10		8	
RW Better	2		0		0		0		0	
(1)+(2)§										
Model Better	7		9		11		12		12	
RW Better	2		0		0		0		0	

Note:

–h denotes forecast horizons for monthly data.

–For each horizon (h), the first column (Cov.) reports empirical coverages given a nominal coverage of 90%. The second column (Leng.) reports the length of forecast intervals in terms of percentage change of the exchange rate. Empirical coverages and lengths are averages across $N(h)$ out-of-sample trials.

–Superscripts a, b, c in the column of Cov. (Leng.) denote rejections of equal coverage accuracy (equal length) in favor of the economic model at a 10%, 5% and 1% confidence level respectively. Superscripts x, y, z are defined analogously for rejections in favor of the random walk.

†–In this panel, a better model is the one with more accurate empirical coverages. RW is the abbreviation of Random Walk.

‡–In this panel, a better model is the one with tighter forecast intervals given equal coverage accuracy.

§–In this panel, a better model is the one with either more accurate coverages or tighter forecast intervals given equal coverage accuracy.

Table 2: Results of Taylor Rule Model Two

	$h=1$		$h=3$		$h=6$		$h=9$		$h=12$	
	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.
Australian Dollar	0.884	7.146^y	0.899	15.086^c	0.918	20.902^a	0.880	26.222^c	0.856	30.677^c
Canadian Dollar	0.825	3.442^c	0.783	6.362^c	0.820	8.815^c	0.846	10.610^c	0.813	13.216^c
Danish Kroner	0.925	8.756^c	0.934	17.791^c	0.959	27.511^z	0.963	32.956	0.947	40.387
French Franc	0.922	8.840^c	0.920	18.740	0.949^c	29.161^c	0.936	34.994^c	0.868	41.330^c
Deutschmark	0.936	9.005	0.879	19.489	0.952	29.658	0.941^a	38.006	0.980	44.355^z
Italian Lira	0.920	9.095^b	0.882	18.558	0.910	27.464^c	0.908	37.325^c	0.921	43.005^c
Japanese Yen	0.915	9.565	0.914	19.752	0.912	29.618	0.937	36.834	0.942	44.455^b
Dutch Guilder	0.908	8.645^c	0.897	18.983^c	0.971	29.391^b	0.990	38.867^z	0.980	46.650^z
Portuguese Escudo	0.916	7.956	0.957	17.924^z	0.909^c	24.196^c	0.889	28.533^z	0.883^a	35.338^c
Swedish Krona	0.861	7.575	0.860	15.679^c	0.851	24.916^c	0.849	31.108^c	0.823	40.458^c
Swiss Franc	0.947	10.008	0.928	20.379^y	0.976	29.578^c	0.988	37.858	0.962	44.675^z
British Pound	0.919	8.614^z	0.933	17.302^c	0.922	26.196^c	0.937	31.371^a	0.957	37.239^c
(1) Coverage Test†										
Model Better	0		0		2		1		1	
RW Better	0		0		0		0		0	
(2) Length Test Given Equal Coverage Accuracy‡										
Model Better	5		6		7		6		7	
RW Better	2		2		1		2		3	
(1)+(2)§										
Model Better	5		6		9		7		8	
RW Better	2		2		1		2		3	

Note:

–h denotes forecast horizons for monthly data.

–For each horizon (h), the first column (Cov.) reports empirical coverages given a nominal coverage of 90%. The second column (Leng.) reports the length of forecast intervals in terms of percentage change of the exchange rate. Empirical coverages and lengths are averages across $N(h)$ out-of-sample trials.

–Superscripts a, b, c in the column of Cov. (Leng.) denote rejections of equal coverage accuracy (equal length) in favor of the economic model at a 10%, 5% and 1% confidence level respectively. Superscripts x, y, z are defined analogously for rejections in favor of the random walk.

†–In this panel, a better model is the one with more accurate empirical coverages. RW is the abbreviation of Random Walk.

‡–In this panel, a better model is the one with tighter forecast intervals given equal coverage accuracy.

§–In this panel, a better model is the one with either more accurate coverages or tighter forecast intervals given equal coverage accuracy.

Table 3: Results of Taylor Rule Model Three

	$h=1$		$h=3$		$h=6$		$h=9$		$h=12$	
	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.
Australian Dollar	0.884	7.163^{y}	0.893	14.925^{c}	0.881	21.067	0.869	25.590^{c}	0.835	28.714^{c}
Canadian Dollar	0.831	3.453^{b}	0.789	6.402^{c}	0.808	8.764^{c}	0.846	10.644^{c}	0.795	12.592^{c}
Danish Kroner	0.925	8.794^{b}	0.924	17.831^{c}	0.959	27.536^{z}	0.963	33.251	0.942	40.036^{a}
French Franc	0.941	8.880^{c}	0.880	18.389^{c}	0.876^{c}	28.548^{b}	0.915	35.443^{a}	0.813	41.350^{c}
Deutschmark	0.945	9.042	0.897	19.642^{z}	0.914	29.677	0.901^{c}	37.291^{a}	0.878	44.520^{y}
Italian Lira	0.906	8.831^{c}	0.868	18.064^{c}	0.902	27.430^{c}	0.877	37.364^{c}	0.803	41.499^{c}
Japanese Yen	0.905	9.181^{c}	0.873	18.910^{c}	0.881	25.700^{c}	0.927	31.259^{c}	0.894	37.049^{c}
Dutch Guilder	0.927	8.910	0.907	19.204^{a}	0.942	29.637	0.951^{a}	36.896^{c}	0.959	46.321^{z}
Portuguese Escudo	0.930	7.961	0.928	16.808^{c}	0.909^{c}	24.059^{b}	0.873	27.868	0.917^{b}	34.980^{c}
Swedish Krona	0.861	7.375^{c}	0.843	15.096^{c}	0.886	24.770^{c}	0.849	31.044^{c}	0.817	38.468^{c}
Swiss Franc	0.965	9.959	0.934	20.433^{y}	0.957	29.418^{c}	0.926^{c}	37.130	0.911	43.546
British Pound	0.919	8.537	0.939	17.397^{b}	0.927	25.809^{c}	0.926	30.749^{c}	0.968	36.825^{c}
(1) Coverage Test†										
Model Better	0		0		2		2		1	
RW Better	0		0		0		0		0	
(2) Length Test Given Equal Coverage Accuracy‡										
Model Better	4		5		5		7		8	
RW Better	0		2		1		0		2	
(1)+(2)§										
Model Better	4		5		7		8		9	
RW Better	0		2		0		0		2	

Note:

–h denotes forecast horizons for monthly data.

–For each horizon (h), the first column (Cov.) reports empirical coverages given a nominal coverage of 90%. The second column (Leng.) reports the length of forecast intervals in terms of percentage change of the exchange rate. Empirical coverages and lengths are averages across $N(h)$ out-of-sample trials.

–Superscripts a, b, c in the column of Cov. (Leng.) denote rejections of equal coverage accuracy (equal length) in favor of the economic model at a 10%, 5% and 1% confidence level respectively. Superscripts x, y, z are defined analogously for rejections in favor of the random walk.

†–In this panel, a better model is the one with more accurate empirical coverages. RW is the abbreviation of Random Walk.

‡–In this panel, a better model is the one with tighter forecast intervals given equal coverage accuracy.

§–In this panel, a better model is the one with either more accurate coverages or tighter forecast intervals given equal coverage accuracy.

Table 4: Results of Taylor Rule Model Four

	$h=1$		$h=3$		$h=6$		$h=9$		$h=12$	
	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.
Australian Dollar	0.895	7.119	0.888	14.409^c	0.902	20.359^c	0.911	24.165^c	0.872	27.492^c
Canadian Dollar	0.814	3.425^c	0.777	6.332^c	0.744	8.490^c	0.769	10.177^c	0.759	12.043^c
Danish Kroner	0.920	8.703^c	0.929	17.534^c	0.964	26.025	0.984^x	30.891^c	0.963	36.443^c
French Franc	0.931	9.065^c	0.860	17.422^c	0.938^c	25.950^c	0.883	30.016^c	0.791	35.192^c
Deutschmark	0.945	8.866^c	0.888	18.839^c	0.894	26.803^c	0.911^c	32.839^c	0.929	38.699^c
Italian Lira	0.891	8.663^c	0.838	17.575^c	0.865	26.307^c	0.777	33.602^c	0.756	38.890^c
Japanese Yen	0.905	9.160^c	0.863	18.708^c	0.861	24.386^c	0.869	28.730^c	0.851	31.697^c
Dutch Guilder	0.936	8.797	0.897	18.368^c	0.914	26.700^c	0.931^c	29.974^c	0.929^c	37.481^c
Portuguese Escudo	0.901	8.183^y	0.884^b	16.237^b	0.909^c	22.354^c	0.905	25.896^c	0.917	30.329^c
Swedish Krona	0.861	7.382^c	0.854	15.095^c	0.869	23.340^c	0.820	30.370^c	0.805	36.487^c
Swiss Franc	0.965	9.644^c	0.940	19.782^a	0.957	27.332^c	0.975	31.004^c	0.956	35.362^c
British Pound	0.904	8.464	0.923	16.287^c	0.854	23.394^c	0.825	27.333^c	0.855	29.796^c
(1) Coverage Test†										
Model Better	0		0		2		2		2	
RW Better	0		0		0		1		0	
(2) Length Test Given Equal Coverage Accuracy‡										
Model Better	8		1		9		9		10	
RW Better	1		0		0		0		0	
(1)+(2)§										
Model Better	8		11		11		11		12	
RW Better	1		0		0		1		0	

Note:

–h denotes forecast horizons for monthly data.

–For each horizon (h), the first column (Cov.) reports empirical coverages given a nominal coverage of 90%. The second column (Leng.) reports the length of forecast intervals in terms of percentage change of the exchange rate. Empirical coverages and lengths are averages across $N(h)$ out-of-sample trials.

–Superscripts a, b, c in the column of Cov. (Leng.) denote rejections of equal coverage accuracy (equal length) in favor of the economic model at a 10%, 5% and 1% confidence level respectively. Superscripts x, y, z are defined analogously for rejections in favor of the random walk.

†–In this panel, a better model is the one with more accurate empirical coverages. RW is the abbreviation of Random Walk.

‡–In this panel, a better model is the one with tighter forecast intervals given equal coverage accuracy.

§–In this panel, a better model is the one with either more accurate coverages or tighter forecast intervals given equal coverage accuracy.

Table 5: Results of Purchasing Power Parity Model

	$h = 1$		$h = 3$		$h = 6$		$h = 9$		$h = 12$	
	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.
Australian Dollar	0.895	7.114^z	0.883	15.558	0.912	21.311	0.880	26.120^c	0.856	30.316^c
Canadian Dollar	0.808	3.513	0.811	6.929	0.814	9.669^b	0.828	11.912^a	0.789	15.303^c
Danish Kroner	0.925	8.735^b	0.929	18.253	0.938	25.891^c	0.969	32.151^b	0.952	38.019^c
French Franc	0.922	8.918^c	0.940	18.137^c	0.979	26.944^c	0.936	31.597^c	0.824	37.655^c
Deutschmark	0.936	9.079	0.935	18.797^c	0.942	27.588^c	1.000	33.585^c	0.990	39.821^c
Italian Lira	0.913	8.794^c	0.875	18.600	0.887	26.619^c	0.954	36.347^c	0.953	42.926^c
Japanese Yen	0.920	9.662^z	0.899	19.903	0.912	28.691^c	0.932	33.973^c	0.899	38.568^c
Dutch Guilder	0.936	8.830	0.935	18.904^c	0.952	27.902^c	1.000	33.468^c	0.990	40.045^c
Portuguese Escudo	0.901	8.049	0.913^a	17.818	0.909^c	23.033^c	0.841	25.579^c	0.900^c	32.050^c
Swedish Krona	0.861	7.541^c	0.876	16.089	0.886	24.345^c	0.855	31.744^a	0.799	37.943^c
Swiss Franc	0.941	9.884^c	0.946	19.709^c	0.988	28.110^c	0.988	33.301^c	0.981	39.887^c
British Pound	0.934	8.643^z	0.939	17.317^c	0.938	25.283^c	0.952	29.418^c	0.930	32.964^c
(1) Coverage Test†										
Model Better	0		1		1		0		1	
RW Better	0		0		0		0		0	
(2) Length Test Given Equal Coverage Accuracy‡										
Model Better	5		5		10		12		11	
RW Better	3		0		0		0		0	
(1)+(2)§										
Model Better	5		6		11		12		12	
RW Better	3		0		0		0		0	

Note:

–h denotes forecast horizons for monthly data.

–For each horizon (h), the first column (Cov.) reports empirical coverages given a nominal coverage of 90%. The second column (Leng.) reports the length of forecast intervals in terms of percentage change of the exchange rate. Empirical coverages and lengths are averages across $N(h)$ out-of-sample trials.

–Superscripts a, b, c in the column of Cov. (Leng.) denote rejections of equal coverage accuracy (equal length) in favor of the economic model at a 10%, 5% and 1% confidence level respectively. Superscripts x, y, z are defined analogously for rejections in favor of the random walk.

†–In this panel, a better model is the one with more accurate empirical coverages. RW is the abbreviation of Random Walk.

‡–In this panel, a better model is the one with tighter forecast intervals given equal coverage accuracy.

§–In this panel, a better model is the one with either more accurate coverages or tighter forecast intervals given equal coverage accuracy.

Table 6: Results of Monetary Model

	$h=1$		$h=3$		$h=6$		$h=9$		$h=12$	
	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.	Cov.	Leng.
Australian Dollar	0.869	7.089	0.853	15.357	0.840	21.353	0.822	25.707^c	0.766	30.099^c
Canadian Dollar	0.791	3.529	0.783	6.854	0.779	9.903	0.787	12.336	0.747	15.176^c
Danish Kroner	0.905	8.771^b	0.893	18.049	0.871	26.413	0.864	31.655^b	0.856	38.564^c
French Franc	0.922	8.830^c	0.910	18.346^c	0.949^b	26.794^c	0.957	32.389^c	0.956	38.113^c
Deutschmark	0.936	8.944^a	0.897	18.615^c	0.875	27.610^c	0.901^c	33.223^c	0.908	40.409^c
Italian Lira	0.913	9.001^c	0.882	18.445^b	0.925	26.613^c	0.954	34.968^c	0.945	41.395^c
Japanese Yen	0.920	9.542	0.919	19.374^c	0.871	28.312^c	0.864	33.401^c	0.814	38.149^c
Dutch Guilder	0.917	8.753^a	0.916	19.384^a	0.962	29.149^b	0.970	38.173	0.908^c	43.896^c
Portuguese Escudo	0.916	8.073	0.957	17.811	0.985	24.971^z	0.968	28.026	1.000	34.598^c
Swedish Krona	0.856	7.460^c	0.837	15.587^c	0.823	22.536^c	0.826	28.641^c	0.781	33.112^c
Swiss Franc	0.929	9.910	0.868	19.539^c	0.793	26.827^c	0.745	31.797^c	0.722	36.189^c
British Pound	0.929	8.398^c	0.928	17.355^c	0.896	25.383^c	0.884	30.600^c	0.850	34.251^c
(1) Coverage Test[†]										
Model Better	0		0		1		1		1	
RW Better	0		0		0		0		0	
(2) Length Test Given Equal Coverage Accuracy[‡]										
Model Better	7		8		8		8		11	
RW Better	0		0		1		0		0	
(1)+(2)[§]										
Model Better	7		8		8		9		12	
RW Better	0		0		1		0		0	

Note:

–h denotes forecast horizons for monthly data.

–For each horizon (h), the first column (Cov.) reports empirical coverages given a nominal coverage of 90%. The second column (Leng.) reports the length of forecast intervals in terms of percentage change of the exchange rate. Empirical coverages and lengths are averages across $N(h)$ out-of-sample trials.

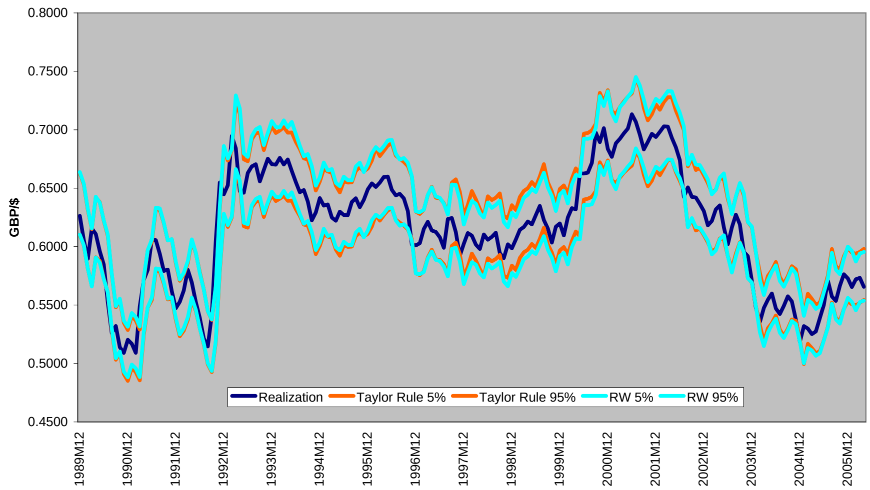
–Superscripts a, b, c in the column of Cov. (Leng.) denote rejections of equal coverage accuracy (equal length) in favor of the economic model at a 10%, 5% and 1% confidence level respectively. Superscripts x, y, z are defined analogously for rejections in favor of the random walk.

†–In this panel, a better model is the one with more accurate empirical coverages. RW is the abbreviation of Random Walk.

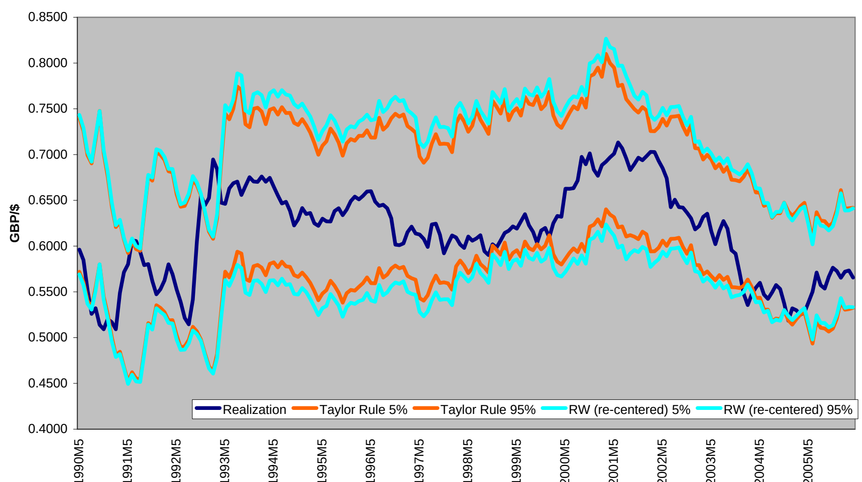
‡–In this panel, a better model is the one with tighter forecast intervals given equal coverage accuracy.

§–In this panel, a better model is the one with either more accurate coverages or tighter forecast intervals given equal coverage accuracy.

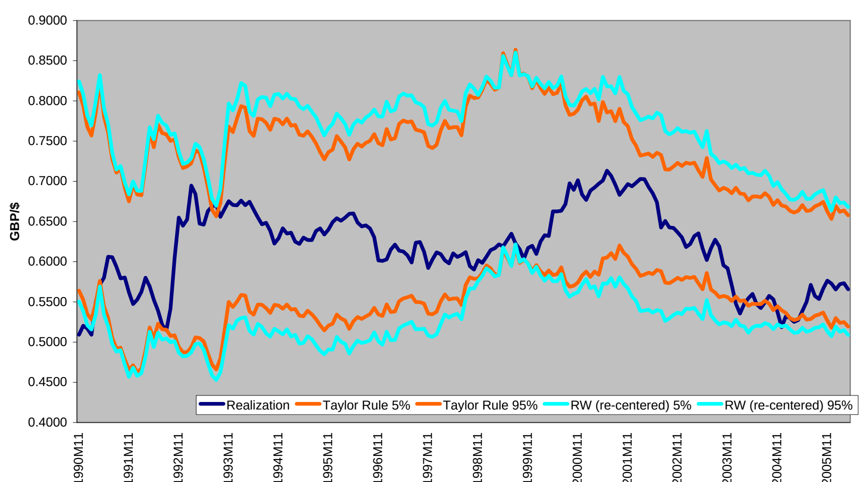
Figure 1: Forecast Intervals of Benchmark Taylor Rule and Random Walk (British Pound)



(a) 1-month-ahead forecast



(b) 6-month-ahead forecast

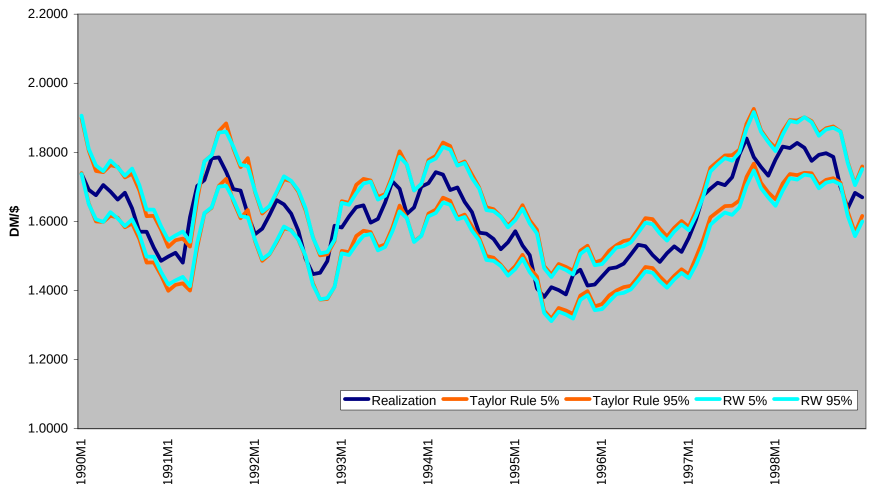


(c) 12-month-ahead forecast

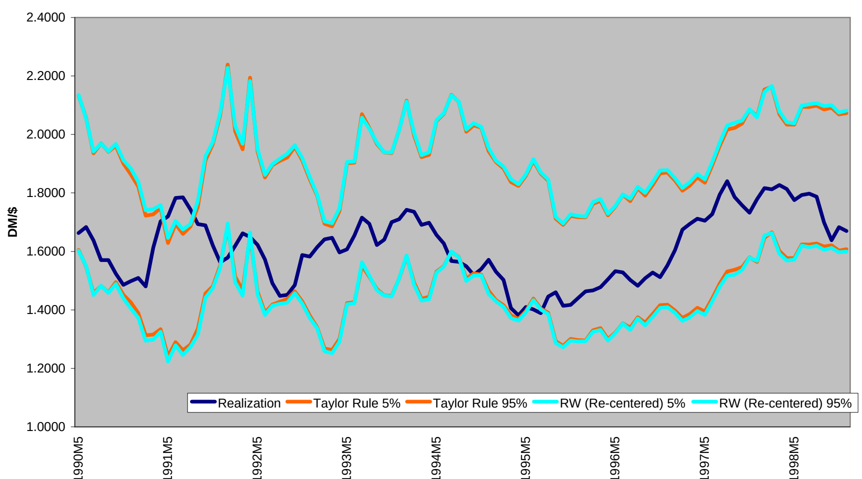
Note:

To facilitate graphical comparisons, the 6- and 12-month-ahead forecast intervals of the random walk have been relocated such that they have the same center as the intervals of the Taylor rule model.

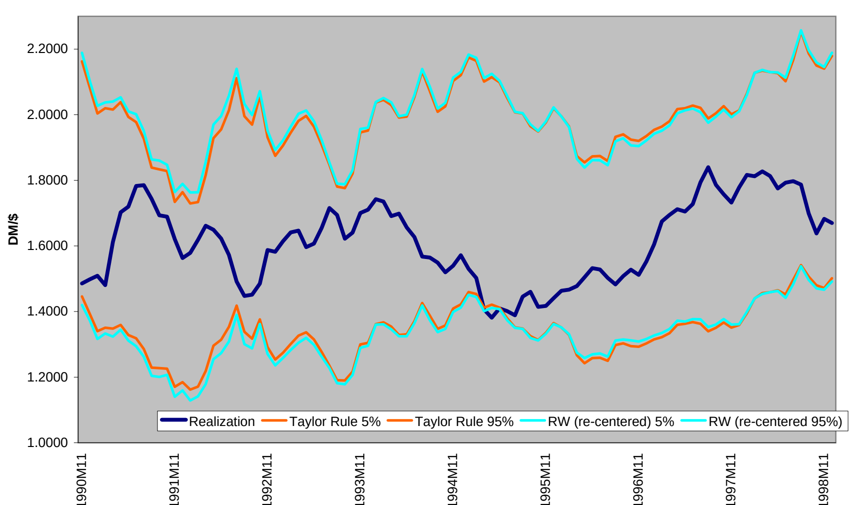
Figure 2: Forecast Intervals of Benchmark Taylor Rule and Random Walk (Deutschmark)



(a) 1-month-ahead forecast



(b) 6-month-ahead forecast



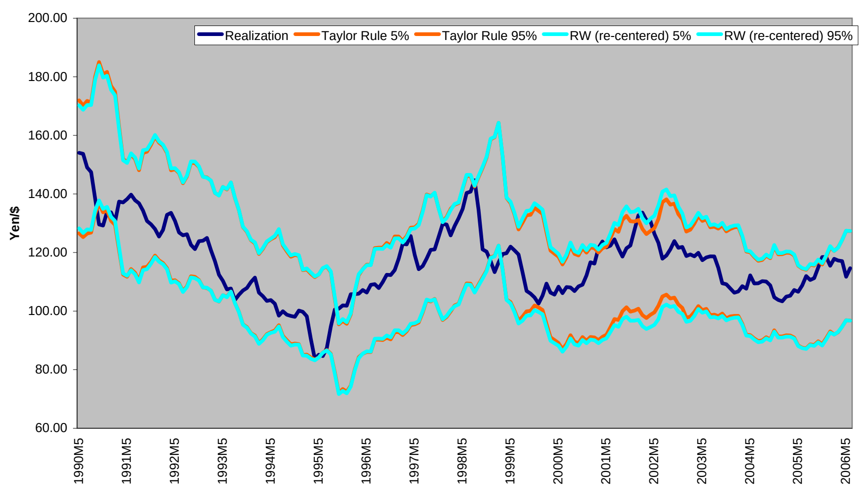
(c) 12-month-ahead forecast

Note:

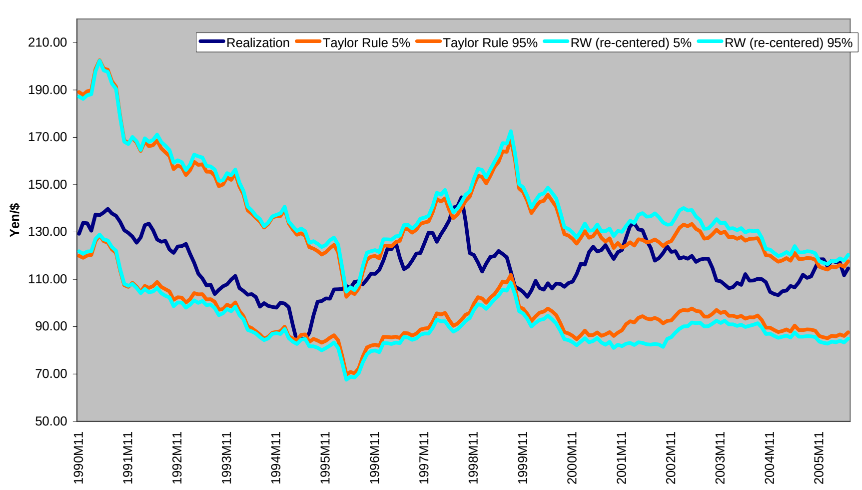
To facilitate graphical comparisons, the 6- and 12-month-ahead forecast intervals of the random walk have been relocated such that they have the same center as the intervals of the Taylor rule model.

Figure 3: Forecast Intervals of Benchmark Taylor Rule and Random Walk (Japanese Yen)

(a) 1-month-ahead forecast



(b) 6-month-ahead forecast



(c) 12-month-ahead forecast

Note:

To facilitate graphical comparisons, the 6- and 12-month-ahead forecast intervals of the random walk have been relocated such that they have the same center as the intervals of the Taylor rule model.

References

[1] **Berkowitz, J. and G. Lorenzo (2001)** "Long-Horizon Exchange Rate Predictability?" *The Review of Economics and Statistics* 83(1), 81-91.

[2] **Boero, G. and E. Marrocu (2004)** "The performance of SETAR models: a regime conditional evaluation of point, interval and density forecasts," *International Journal of Forecasting* 20(2), 305-320.

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

[4] **Chinn, M. D. and R. A. Meese (1995)** "Banking on Currency Forecasts: How Predictable Is Change in Money?" *Journal of International Economics* 38, 161-178.

[5] **Christoffersen, P. F. (1998)** "Evaluating Interval Forecasts," *International Economic Review* 39(4), 841-862.

[6] **Clarida, R., Gali, J. and M. Gertler (1998)** "Monetary Rules in Practice: Some International Evidence," *European Economic Review* 42, 1033-1067.

[7] **Clarida, R.H., L. Sarno, M. P. Taylor, and G. Valente (2003)** "The Out-of-Sample Success of Term Structure Models as Exchange Rate Predictors: A Step Beyond," *Journal of International Economics* 60, 61-83.

[8] **Clark, T. and K. D. West (2006)** "Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis," *Journal of Econometrics* 135 (1-2), 155-186.

[9] **Clark, T. and K. D. West (2007)** "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models," *Journal of Econometrics* 138(1), 291-311.

[10] **Corradi, V. and N. Swanson (2006a)** "Predictive Density and Conditional Confidence Interval Accuracy Tests," *Journal of Econometrics* 135, 187-228.

[11] **Corradi, V. and N. Swanson (2006b)** "Predictive Density Evaluation," In Grangerm C.W.J., Elliott, G., Timmerman, A. (Eds.), *Handbook of Economic Forecasting*, Elsevier, Amsterdam.

[12] **Diebold, F. X., Gunther, T. A. and A. S. Tay (1998)** "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review* 39(4), 863-883.

[13] **Diebold, F. X., Hahn, J. and A .S. Tay (1999)** "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns of Foreign Exchange," *Review of Economics and Statistics* 81, 661-673.

[14] **Diebold, F. X. and R. S. Mariano (1995)** "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics* 13, 253-263.

[15] **Engel, C., N. Mark, and K. West (2007)** "Exchange Rate Models are Not as Bad as You Think," in NBER Macroeconomics Annual.

[16] **Engel, C. and K. West (2005)** "Exchange Rates and Fundamentals," *Journal of Political Economy* 113, 485-517.

[17] **Engel, C. and K. D. West (2006)** "Taylor Rules and the Deutschmark-Dollar Real Exchange Rate," *Journal of Money, Credit and Banking* 38(5), 1175-1194.

[18] **Engel, C., J. Wang, and J. J. Wu (2008)** "Can Long-horizon Data Beat the Random Walk under the Engel-West Explanation?" Working Paper, University of Wisconsin-Madison, the Federal Reserve Bank of Dallas, and the Federal Reserve Board.

[19] **Faust, J., J. Rogers, and J. Wright, (2003)** "Exchange Rate Forecasting: The Errors Weve Really Made," *Journal of International Economics* 60, 35-59.

[20] **Giacomini, R. and H. White (2006)** "Tests of Conditional Predictive Ability," *Econometrica* 74, 1545-1578.

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

[22] **Hall, P., R.C.L. Wolff, and Q. Yao (1999)** "Methods for Estimating a Conditional Distribution Function," *Journal of American Statistical Association* 94, 154-163.

[23] **Hong, Y., H. Li, and F. Zhao (2007)** "Can the Random Walk be Beaten in Out-of-Sample Density Forecasts: Evidence from Intraday Foreign Exchange Rates," *Journal of Econometrics* 141, 736-776.

[24] **J.P. Morgan (1996)** RiskMetrics Technical Document.

[25] **Kilian, L. (1999)** "Exchange Rates And Monetary Fundamentals: What Do We Learn From Long-Horizon Regressions?" *Journal of Applied Econometrics* 14, 491-510.

[26] **Kilian, L. and M. Taylor (2003)** "Why Is It So Difficult to Beat the Random Walk Forecast of Exchange Rate?" *Journal of International Economics* 60, 85-107.

[27] **MacDonald, R. and M. P. Taylor (1994)** "The Monetary Model of the Exchange Rate: Long-Run Relationships, Short-Run Dynamics, and How to Beat a Random Walk," *Journal of International Money and Finance* 13, 276-290.

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

[29] **Mark, N. C. (2007)** "Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dynamics," *Manuscript*, University of Notre Dame.

[30] **Mark, N. and D. Sul (2001)** "Nominal Exchange Rates and Monetary Fundamentals Evidence from a Small Post-Bretton Woods Panel," *Journal of International Economics* 53, 29-52.

[31] **Meese, R. A. and K. Rogoff (1983)** "Empirical Exchange Rate Models of the Seventies: Do they Fit Out-of-Sample?" *Journal of International Economics* 14, 3-24.

[32] **Molodtsova, T. and D. H. Papell (2008)** "Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals," *Manuscript*, University of Houston.

[33] **Molodtsova, T., A. Nikolsko-Rzhevskyy, and D. H. Papell (2008)**, "Taylor Rules with Real-Time Data: A Tale of Two Countries and One Exchange Rate," *Journal of Monetary Economics* 55, S63-S79.

[34] **Nason, J. and J. Rogers (2008)** "Exchange Rates and Fundamentals: A Generalization," International Finance Discussion Papers, Number 948, Board of Governors of the Federal Reserve System.

[35] **Newey, W. K. and K. D. West (1987)** "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix" *Econometrica* 55, 703-708.

[36] **Rossi, B. (2005)** "Testing Long-horizon Predictive Ability with High Persistence, and the Meese-Rogoff Puzzle," *International Economic Review* 46(1), 61-92.

[37] **Sarno, L. and G. Valente (2005)** "Empirical Exchange Rate Models and Currency Risk: Some Evidence from Density Forecasts," *Journal of International, Money and Finance* 24(2), 363-385.

[38] **West, K. D. (1996)** "Asymptotic Inference About Predictive Ability," *Econometrica* 64, 1067-1084.

[39] **Wu, J. J. (2007)** "Robust Semiparametric Forecast Intervals," *Manuscript*, the Federal Reserve Board.

APPENDIX

A.1 Monetary and Taylor Rule Models

In this section, we describe the monetary and Taylor rule models used in the paper.

A.1.1 Monetary Model

Assume the money market clearing condition in the home country is

$$m_t = p_t + \gamma y_t - \alpha i_t + v_t,$$

where m_t is the log of money supply, p_t is the log of aggregate price, i_t is the nominal interest rate, y_t is the log of output, and v_t is money demand shock. A symmetric condition holds in the foreign country and we use an asterisk in subscript to denote variables in the foreign country. Subtracting foreign money market clearing condition from the home, we have

$$i_t - i_t^* = \frac{1}{\alpha}\left[-(m_t - m_t^*) + (p_t - p_t^*) + \gamma(y_t - y_t^*) + (v_t - v_t^*)\right]. \tag{A.1.1}$$

The nominal exchange rate is equal to its purchasing power value plus the real exchange rate:

$$s_t = p_t - p_t^* + q_t. \tag{A.1.2}$$

The uncovered interest rate parity in financial market takes the form

$$E_t s_{t+1} - s_t = i_t - i_t^* + \rho_t, \tag{A.1.3}$$

where ρ_t is the uncovered interest rate parity shock. Substituting equations (A.1.1) and (A.1.2) into (A.1.3), we have

$$s_t = (1-b)\left[m_t - m_t^* - \gamma(y_t - y_t^*) + q_t - (v_t - v_t^*)\right] - b\rho_t + bE_t s_{t+1}, \tag{A.1.4}$$

where $b = \alpha/(1+\alpha)$. Solving s_t recursively and applying "no-bubbles" condition, we have

$$s_t = E_t \left\{ (1-b) \sum_{j=0}^{\infty} b^j \left[m_{t+j} - m^*_{t+1} - \gamma(y_{t+j} - y^*_{t+1}) + q_{t+1} - (v_{t+j} - v^*_{t+j}) \right] - b \sum_{j=1}^{\infty} b^j \rho_{t+j} \right\}. \quad \text{(A.1.5)}$$

In the standard monetary model, such as Mark (1995), purchasing power parity ($q_t = 0$) and uncovered interest rate parity hold ($\rho_t = 0$). Furthermore, it is assumed that the money demand shock is zero ($v_t = v^*_t = 0$) and $\gamma = 1$. Equation (A.1.5) reduces to

$$s_t = E_t \left\{ (1-b) \sum_{j=0}^{\infty} b^j \left(m_{t+j} - m^*_{t+j} - (y_{t+j} - y^*_{t+j}) \right) \right\}.$$

A.1.2 Taylor Rule Model

We follow Engel and West (2005) to assume that both countries follow the Taylor rule and the foreign country targets the exchange rate in its Taylor rule. The interest rate differential is

$$i_t - i^*_t = \delta_s(s_t - \bar{s}^*_t) + \delta_y(y^{gap}_t - y^{gap*}_t) + \delta_\pi(\pi_t - \pi^*_t) + v_t - v^*_t, \quad \text{(A.1.6)}$$

where $\bar{s}^*_t$ is the targeted exchange rate. Assume that monetary authorities target the PPP level of the exchange rate: $\bar{s}^*_t = p_t - p^*_t$. Substituting this condition and the interest rate differential to the UIP condition, we have

$$s_t = (1-b)(p_t - p^*_t) - b\left[\delta_y(y^{gap}_t - y^{gap*}_t) + \delta_\pi(\pi_t - \pi^*_t) + v_t - v^*_t\right] - b\rho_t + bE_t s_{t+1}, \quad \text{(A.1.7)}$$

where $b = \frac{1}{1+\delta_s}$. Assuming that uncovered interest rate parity hold ($\rho_t = 0$) and monetary shocks are zero, equation (A.1.7) reduces to the benchmark Taylor rule model in our paper:

$$s_t = E_t \left\{ (1-b) \sum_{j=0}^{\infty} b^j (p_{t+j} - p^*_{t+j}) - b \sum_{j=0}^{\infty} b^j (\delta_y(y^{gap}_{t+j} - y^{gap*}_{t+j}) + \delta_\pi(\pi_{t+j} - \pi^*_{t+j})) \right\}.$$

A.2 Long-horizon Regressions

In this section, we derive long-horizon regressions for the monetary model and the benchmark Taylor rule model.

A.2.1 Monetary Model

In the monetary model,

$$s_t = E_t \left\{ (1-b) \sum_{j=0}^{\infty} b^j \left(m_{t+j} - m^*_{t+j} - (y_{t+j} - y^*_{t+j}) \right) \right\},$$

where m_t and y_t are logarithms of domestic money stock and output, respectively. The superscript $*$ denotes the foreign country. Money supplies (m_t and m^*_t) and total outputs (y_t and y^*_t) are usually I(1) variables.

The general form considered in Engel, Wang, and Wu(2008) is:

$$\begin{aligned} s_t &= (1-b) \sum_{j=0}^{\infty} b^j E_t \alpha^{'} \mathbf{D}_t \\ (I_n - \Phi(L))\Delta \mathbf{D}_t &= \varepsilon_t \\ E(\varepsilon_{t+j} | \varepsilon_t, \varepsilon_{t-1}, ...) &\equiv E_t(\varepsilon_{t+j}) = 0, \forall j \geq 1, \end{aligned} \tag{A.2.1}$$

where n is the dimension of $\mathbf{D}_t$ and I_n is an $n \times n$ identity matrix. L is the lag operator and $\Phi(L) = \phi_1 L + \phi_2 L^2 + ... + \phi_p L^p$. Assume $\Phi(1)$ is non-diagonal and the covariance matrix of ε_t is given by $\Omega = E_t[\varepsilon_t \varepsilon_t']$. We assume that the change of fundamentals follows a VAR(p) process in our setup. From proposition 1 of Engel, Wang, Wu (2008), we know that for a fixed discount factor b and $p \geq 2$,

$$s_{t+h} - s_t = \beta_h z_t + \delta^{'}_{0,h} \Delta \mathbf{D}_t + ... + \delta^{'}_{p-2,h} \Delta \mathbf{D}_{t-p+2} + \zeta_{t+h}$$

is a correctly specified regression where the regressors and errors do not correlate. In the case of $p = 1$, the long-horizon regressions reduces to

$$s_{t+h} - s_t = \beta_h z_t + \zeta_{t+h}.$$

Following the literature, for instance Mark (1995), we do not include $\Delta \mathbf{D}_t$ and its lags in our long-horizon regressions. The monetary model can be written in the form of (A.2.1) by setting $\mathbf{D}_t = [m_t \quad m^*_t \quad y_t \quad y^*_t]'$, $\alpha = [1 \quad -1 \quad -1 \quad 1]'$. By definition, $z_t = s_t - (m_t - m^*_t) + (y_t - y^*_t)$. This corresponds to $\beta_{m,h} = 1$, $\mathbf{X}_{m,t} = s_t - (m_t - m^*_t) + (y_t - y^*_t)$ in equation (1) of section 3.

A.2.2 Taylor Rule Model

In the Taylor rule model,

$$s_t = E_t \left\{ (1-b) \sum_{j=0}^{\infty} b^j (p_{t+j} - p^*_{t+j}) - b \sum_{j=0}^{\infty} b^j (\delta_y (y^{gap}_{t+j} - y^{gap*}_{t+j}) + \delta_\pi (\pi_{t+j} - \pi^*_{t+j})) \right\},$$

where p_t, y^{gap}_t and π_t are domestic aggregate price, output gap and inflation rate, respectively. δ_y and δ_π are coefficients of the Taylor rule model. The aggregate prices p_t and p^*_t are usually I(1) variables. Inflation and output gap are more likely to be I(0). Engel, Wang, and Wu (2008) consider a setup which includes both stationary and non-stationary variables:

$$\begin{aligned} s_t = (1-b) \sum_{j=0}^{\infty} b^j E_t \left[f_{1t+j} \right] + b \sum_{j=0}^{\infty} b^j E_t \left[f_{2t+j} + u_{2t+j} \right] \\ f_{1t} &= \alpha_1' \mathbf{D_t} \sim I(1) \\ f_{2t} &= \alpha_2' \mathbf{\Delta D_t} \sim I(0) \\ u_{2t} &= \alpha_3' \mathbf{\Delta D_t} \sim I(0) \\ (I_n - \Phi(L)) \Delta \mathbf{X}_t &= \varepsilon_t, \end{aligned} \tag{A.2.2}$$

where f_{1t} and f_{2t} (u_{2t}) are observable (unobservable) fundamentals. $\Delta \mathbf{D}_t$ is the first difference of $\mathbf{D}_t$, which contains I(1) economic variables.[17]

From proposition 2 of Engel, Wang, and Wu (2008), we know that for a fixed discount factor b and $h \geq 2$,

$$s_{t+h} - s_t = \tilde{\beta}_h z_t + \sum_{k=0}^{p-1} \tilde{\delta}'_{k,h} \Delta \mathbf{D}_{t-k} + \tilde{\zeta}_{t+h} \tag{A.2.3}$$

is a correctly specified regression, where the regressors and errors do not correlate. In the case of $p = 1$, the long-horizon regressions reduces to

$$s_{t+h} - s_t = \tilde{\beta}_h z_t + \tilde{\zeta}_{t+h}.$$

[17] To incorporate I(0) economic variables, $\mathbf{D}_t$ contains the levels of I(1) variables and the summation of I(0) variables from negative infinity to time t.

The Taylor rule model can be written into the form of (A.2.2) by setting

$$\mathbf{D}_t = \left[p_t \quad p_t^* \quad \sum_{s=-\infty}^{t} y_s^{gap} \quad \sum_{s=-\infty}^{t} y_s^{gap*}, \sum_{s=-\infty}^{t} \pi_s \quad \sum_{s=-\infty}^{t} \pi_s^* \right]'.$$

By definition, $z_t = s_t - p_t + p_t^* + \frac{b}{1-b}(\delta_y(y_t^{gap} - y_t^{gap*}) + \delta_\pi(\pi_t - \pi_t^*))$. This corresponds to $\beta_{m,h} = [1 \quad \frac{b}{1-b}\delta_y \quad \frac{b}{1-b}\delta_\pi]$ and $\mathbf{X}_{m,t} = [q_t \quad y_t^{gap} - y_t^{gap*} \quad \pi_t - \pi_t^*]$, where $q_t = s_t - p_t + p_t^*$ is the real exchange rate. $\beta_{m,h}$ and $\mathbf{X}_{m,t}$ can be defined differently. For instance, $\beta_{m,h} = 1$ and $\mathbf{X}_{m,t} = s_t - p_t + p_t^* + \frac{b}{1-b}(\delta_y(y_t^{gap} - y_t^{gap*}) + \delta_\pi(\pi_t - \pi_t^*))$. Our results do not change qualitatively under this alternative setup.