Forecasting Exchange Rates Out-of-Sample with Panel Methods and Real-Time Data†

Onur Ince*

Abstract

This paper evaluates out-of-sample exchange rate forecasting with Purchasing Power Parity (PPP) and Taylor rule fundamentals for 9 OECD countries vis-à-vis the U.S. dollar over the period from 1973:Q1 to 2009:Q1 at short and long horizons. In contrast with previous work, which reports "forecasts" using revised data, I construct a quarterly real-time dataset that incorporates only the information available to market participants when the forecasts were made. Using bootstrapped out-of-sample test statistics, the exchange rate model with Taylor rule fundamentals performs better at the one-quarter horizon and panel estimation is not able to improve its performance. The PPP model, however, forecasts better at the 16-quarter horizon and its performance increases in panel framework. The results are in accord with previous research on PPP and Taylor rule models.

Keywords: Exchange Rate Forecasting, Taylor Rules, Real-Time Data, Out-of-Sample Test Statistics

JEL Classification: C23, C53, E32, E52, F31, F47

---

† I thank David Papell for encouragement and guidance, and Chris Murray, Sebnem Kalemli-Ozcan, Nelson Mark, Lutz Kilian, Dean Croushore, Tanya Molodtsova, James Morley, Claude Lopez, Mark Strazicich and Vania Stavrakeva for helpful comments and discussions.

* Department of Economics, Appalachian State University, Boone, NC 28608-2037. Tel: +1 (828) 262-4033. Email: inceo@appstate.edu

1. Introduction

Following the collapse of the Bretton-Woods system, the introduction of flexible exchange rate regimes attracted much attention to the area of international macroeconomics in an attempt to explain exchange rate behavior. Theoretical papers such as Dornbusch (1976), which extended the Mundell-Fleming model to incorporate rational expectations and sticky prices and introduced overshooting as an explanation for high exchange rate variability, and empirical work such as Frankel (1979), which found success in estimating empirical exchange rate models, inspired research in this field by pointing out the ability of macroeconomic models to explain exchange rate variability.

The seminal papers by Meese and Rogoff (1983a, 1983b) put an end to the atmosphere of optimism in exchange rate economics by concluding that empirical exchange rate models do not perform better than a random walk model out-of-sample. Their finding is still hard to overturn more than two decades later. Cheung, Chinn, and Pascual (2005), for example, examine out-of-sample performance of the interest rate parity, monetary, productivity-based and behavioral exchange rate models and conclude that none of the models consistently outperforms the random walk at any horizon.

Are empirical exchange rate models really as bad as we think? Recent studies have found evidence of exchange rate predictability using either panels or innovative modeling approaches. Engel, Mark, and West (2008) use panel specifications of the monetary, Purchasing Power Parity (PPP) and Taylor (1993) rule models, Rossi (2006) uses the monetary model in the presence of a structural break, Gourinchas and Rey (2007) use an external balance model, Molodtsova and Papell (2009) use a heterogeneous symmetric Taylor rule model with smoothing, and Cerra and Saxena (2010) use a broad panel specification of the monetary model.

A common problem with the papers discussed above is their reliance on ex-post revised data for the forecasting analysis. Macroeconomic data are updated when new data become available and frequently revised over time. These revisions can be substantial and were not available to either policymakers or market participants at the time forecasts were made. Therefore, out-of-sample forecast evaluations based on ex-post revised data yield misleading inference about the exchange rate models, and information problems of market agents are not accounted in the analysis. As Rossi (2005) emphasizes, to forecast economic variables which are driven by persistent and permanent shocks, the econometrician might measure agent's probability distribution poorly by using actual realized values of future explanatory variables. To forecast exchange rates, which are primarily driven by expectations, real-time data would be more advantageous due to capturing the information set of market participants as closely as possible in contrast to ex-post revised data and actual realized values of future explanatory variables.

Out-of-sample forecasts of exchange rate models may be influenced by data revisions in many different ways. First, estimated parameters of the candidate models will vary because the data used for in-sample estimation is different. Changes in the parameter estimates could be striking if the forecasting model contains a latent variable whose value is subject to variation due to data revisions, such as output gap in Taylor rule models. Second, changes in parameter estimates induce candidate models to produce different one- and multi-step ahead out-of-sample forecasts. Consequently, out-of-sample inferences based on forecast errors may suggest selecting a different model. Third, due to differences in timing and magnitudes of data revisions across countries, model specifications themselves can be subject to change. More specifically, forecasts generated with time-series regressions in real-time could dominate panel specifications when the level of heterogeneity, arises from differences in data revisions across countries, is high. Although all of the abovementioned reasons suggest that out-of-sample predictive ability of exchange rate models should be evaluated using real-time data, it is still very rare in the exchange rate literature.

The first paper to use real-time data to evaluate nominal exchange rate predictability is Faust, Rogers and Wright (2003). Examining the predictive ability of Mark's (1995) monetary model using real-time data for Japan, Germany, Switzerland and Canada vis-à-vis the U.S, they report that the models consistently perform better using real-time data than fully revised data. However, none of the models perform better than the random walk model. More recently, Molodtsova, Nikolsko-Rzhevskyy, and Papell (2008, 2011) find evidence of predictability with Taylor rule fundamentals using real-time data for the Deutschmark/dollar and Euro/dollar exchange rates. Molodtsova, Nikolsko-Rzhevskyy, and Papell (2008) find evidence of out-of-sample predictability with Taylor rule fundamentals only using real-time data as opposed to ex-post revised data and confirm the conclusion of Faust, Rogers and Wright (2003) that exchange rate dynamics might react more to the market's contemporaneous beliefs about the fundamentals than true underlying fundamentals.

There are no studies on exchange rate forecasting with real-time data for a reasonably large number of countries over the post Bretton Woods period because of the limited availability of real-time data for countries other than the U.S. In this paper, I construct a quarterly real-time dataset that contains 9 OECD countries (Australia, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, the United Kingdom) vis-à-vis the U.S. dollar over the period from 1973:Q1 to 2009:Q1 to evaluate both short and long-horizon out-of-sample forecasting performance of the linear exchange models using PPP and Taylor rule fundamentals. I construct real-time price and inflation data from the International Financial Statistics (IFS) country pages using the consumer price index (CPI), and estimate real-time output gaps using the industrial production index.

A problem associated with recent papers presenting evidence of exchange rate predictability is that these studies employ only a test developed by Clark and West (2006) (henceforth, CW test). The CW statistic adjusts the Diebold and Mariano (1995) and West (1996) (henceforth, DMW test) statistic to correct for size distortions. If two models are non-nested, the DMW test is appropriate to compare the mean square prediction errors (MSPE's). Applying DMW test to compare the MSPE's of two nested models, however, leads to non-normal test statistics, and using standard normal critical values usually results in very poorly sized tests with far too few rejections of the null. This is a problem for out-of-sample exchange rate forecasting because, since the null is a random walk, all tests with structural models are nested. While the CW adjustment produces a test with correct size, Rogoff and Stavrakeva (2008) argue that it cannot evaluate forecasting performance because it does not test the null hypothesis of equal MSPE's of the random walk and the structural model. In order to satisfy the conditions for a "good" exchange rate forecasting model, empirical studies need to present evidence that the exchange rate model has MSPE that is significantly smaller than that of the random walk model, which cannot be done solely with CW test in the case of forecasting bias.[1] They advocate the use of DMW tests with bootstrapped critical values to produce correctly sized tests.

Engel, Mark, and West (2008) find that panel error-correction exchange rate models with PPP fundamentals are able to produce large improvements in out-of-sample forecasting at longer horizons.[2] Because they use ex-post revised data, the exchange rate models in their study contain future information that was not available to market participants. "Forecasts" that are produced using future news in the information set of the linear model cannot be evaluated within an out-of-sample

[1] Rogoff and Stavrakeva (2008) consider the scale bias where the observed value is over- or under predicted by a certain percent.

[2] Engel, Mark, and West (2008) use monetary and Taylor Rule models as well. However, the out-of-sample predictability of the PPP model dominates the other two models at longer horizons.

forecasting exercise. Forecasts with real-time data, however, do not contain any unrealized future information in the information set of the linear model, and thus are a true out-of-sample forecast.

Molodtsova and Papell (2009) find evidence of out-of-sample predictability with the Taylor model at short horizon using single-equation estimation. Although they use ex-post revised data to calculate inflation, they estimate output gaps with quasi-real-time data in order to capture the information available to central banks as closely as possible. Quasi-real-time data is constructed with ex-post revised data, but the trends do not contain future observations and the data points are used with a lag for estimation. While quasi-real-time data does not contain future observations, it captures revisions which are not available to market participants. Therefore, forecasting exercises with quasi-real time data are also not true out-of-sample forecasts.

This paper reevaluates out-of-sample predictive ability of PPP and Taylor rule-based exchange rate models that might have produced fragile conclusions with revised data in the earlier studies. Using a newly constructed real-time dataset for 9 OECD countries vis-à-vis the U.S. dollar, out-of-sample forecasting power of PPP and Taylor rule models are investigated within single-equation and panel frameworks based on bootstrapped DMW and CW test statistics.[3] The out-of-sample forecasting results with PPP fundamentals confirm the findings in Engel, Mark, and West (2008) that the predictability of the PPP model increases with the panel specification and the PPP model has higher predictive power at long horizons. Evidence of long-term predictability with the PPP model is found for 7 out of 9 countries with the CW test and 5 out of 9 countries with the DMW test against the driftless random walk. The exchange rate model with PPP fundamentals using panel data outperforms the random walk with drift for all the countries in the sample at the 16-quarter horizon regardless of which test statistic is used.

[3] I would like to evaluate the out-of-sample forecasting ability of the monetary model. However, it is not possible to obtain a coherent series of real-time money supply for all the countries.

The predictability with Taylor rule fundamentals, in contrast, is greatest with the single-equation specification, and the Taylor rule model has higher forecasting power at the short horizon as indicated in Molodtsova and Papell (2009). Evidence of short-term predictability with Taylor rule model is found for 1 out of 9 countries with both test statistics against the driftless random walk with single-equation estimation. The exchange rate model with Taylor rule fundamentals using a single-equation framework outperforms the random walk with drift for 3 out of 9 countries with the CW test and 5 out of 9 countries with the DMW test at the one-quarter horizon.

The results are in accord with previous research on PPP and Taylor rule models. The PPP model works best with the panel specification at the 16-quarter horizon. Research on PPP shows no evidence of short-run PPP, and Papell (1997) finds considerably more support for long-run PPP with panel methods than with univariate tests. Since the persistence of deviations from PPP is relatively homogeneous across countries, panels help to reduce the noise and increase the forecasting power of the PPP model.

Out-of-sample forecasts that are based on pooled Taylor rules with fixed effects, however, are unable to outperform time-series regression forecasts. A number of studies have found evidence that monetary policy rules implemented by central banks are subject to substantial heterogeneity. For example, Clarida, Gali and Gertler (1998) provide empirical evidence of how interest rate reaction functions vary among OECD countries. Gerdesmeier, Mongelli and Roffia (2007) compare the monetary policies implemented by the Eurosystem, the Fed and the Bank of Japan, and find considerable differences. Thus, weak forecasting performance of Taylor rules in panel specification may reflect heterogeneity in adjustment to equilibrium. As Mark and Sul (2012) emphasize, when the heterogeneity is great, panels do not generate more accurate forecasts than time-series regressions. Since central banks target short-term nominal interest rates as the instrument of monetary policy, higher forecasting power of the Taylor rule model in a single-equation framework at the short horizon is plausible.

2. Data

The real-time quarterly data used in this study covers the post-Bretton Woods period from 1973:Q1 to 2009:Q1 for 10 OECD countries: Australia, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, the United Kingdom, and the United States. The dataset is constructed from the country tables of IMF's *International Financial Statistics* (IFS) books, regularly published on a monthly basis since 1948. The real-time data has the usual triangular format with vintage dates on the horizontal axis and calendar dates for each observation on the vertical axis. The term vintage corresponds to the date when a time series of data becomes available to market participants. There is typically a one-quarter lag between the vintage date and the latest data point at that vintage. The real-time data at time t actually represents data through period t-1. For each subsequent quarter, the new vintage includes both newly released data and revisions to the historical data. The first vintage in the real-time dataset is for 1973:Q1 and the data series in each vintage start from 1958:Q1.

Seasonally adjusted industrial production index (IFS line 66c) is used as a measure of countries' income, since quarterly GDP data are not consistently published and not available for some countries for much of the time span. The price level in the economy is measured by the consumer price index (CPI) (IFS line 64) and seasonally adjusted by applying a one-sided moving average of the current observation and 3-lagged values. The inflation rate is the annual inflation rate calculated using the CPI over the previous 4 quarters.

The output gap is calculated as the percentage deviation of actual output from a Hodrick-Prescott (1997) (HP) trend.[4] For the first vintage, the trend is calculated using the data for 1958:Q1-1972:Q4, for the second vintage, it is calculated using the data for 1958:Q1-1973:Q1, and so on. As with any method that uses a one-sided filter, the estimations might be subject to end-of-sample uncertainty which is exacerbated with real-time data, consisting of the last observations in each data vintage. To take into account the end-of-sample uncertainty in output gap estimation using real-time data, I follow Watson's (2007) method using an AR (8) model to forecast the output growth 12-quarters ahead before calculating the trend.[5] [6]

The release dates for real-time variables vary across countries and the timing of data release is very crucial for forecast evaluation. For example, the industrial production index for Germany is released approximately 38 days after the end of the reference month, while the U.S. industrial production index is released from 12 to 18 days after the reference month. To minimize the time between the release of the data and the start of the forecast, the quarterly real-time dataset is constructed using the data available in second month of each quarter. Nominal exchange rates are taken from the IFS CD-ROM (IFS line ae) defined as the end-of-period U.S. dollar price of a unit of foreign currency.[7] Exchange rates for the Euro area after 1998 are normalized by fixing foreign currency per dollar to the Euro/Dollar rate as in Engel, Mark, and West (2008).

---

[4] The smoothness parameter for HP filter is 1600 with quarterly data.

[5] While Watson (2007) also suggests to backcast the series, the series in each data vintage extends through 1958:Q1, which is long enough to remove the distortions in the beginning of the sample created by a one-sided filter.

[6] HP Filter is selected as the most commonly used filter in the literature. Ince and Papell (2013) also provide the evidence that correlations between real-time and ex-post output gap estimates with different filters for the same countries are similar.

[7] Since quarterly averaged exchange rates might cause serial correlation for exchange rate changes, I use the end-of-period exchange rates.

The series of real-time inflation and output gaps are constructed from the diagonal elements of the real-time data matrix and contain only the latest available observations at each period. For each country, this data represents a vector of quarterly observations from 1973:Q1 to 2009:Q1, thus resulting in 145 observations.

Table 1 presents summary statistics for real-time and revised inflation and output gap for each country in the sample. Two observations are apparent: First, the differences between average real-time and revised inflation rates are very close as opposed to the differences between average real-time and revised output gaps. The differences between the average real-time and revised inflation varies from 0.001 percentage points (for Japan) to 0.076 percentage points (for U.K.), while the difference between the average real-time and revised output gap varies from 0.725 percentage points (for Sweden) to 2.134 percentage points (for Italy). Second, average real-time output gaps are negative for all the countries, which implies that the output gaps are being revised upwards on average. According to the summary statics in Table 1, policy recommendations based on real-time and revised data may differ substantially with most of the differences coming from the revisions in output gaps.

3. Methodology

The econometric analysis in this study is based on panel estimation of the predictive regression,

$$s_{it+k} - s_{it} = \beta_k z_{it} + \varepsilon_{it+k} \quad (1)$$

where $z_{it} = f_{it} - s_{it}$ and $\varepsilon_{it} = \zeta_i + \theta_t + u_{it}$.[8] In the predictive regression, s_{it} denotes the natural log of the nominal exchange rate, measured as the domestic price of U.S. dollar (which serves as base currency) for country i at time t. The deviation of the exchange rate from its equilibrium value is denoted by z, and f stands for the fundamental in the exchange rate model that is determined either by PPP or Taylor rule. The forecast horizon k, takes on the value of 1 for short-horizon and 16 for long-horizon regressions. The regression error, ε_{it}, contains unobserved components, where ζ_i is the individual specific effect, θ_t is the time-specific effect, and u_{it} is the residual idiosyncratic error.

3.1 PPP Fundamentals

Numerous studies that test for unit roots in real exchange rates using panels of industrialized countries have found strong rejections in the post-1973 period. The strong rejections of unit roots encourage testing the forecasting power of exchange rate models with PPP fundamentals. Recently, Engel, Mark, and West (2008) have shown that PPP fundamentals forecast well at long horizons. Rogoff and Stavrakeva (2008) also conclude that PPP specification performs the best out of all the specifications they try.[9]

Under PPP fundamentals,

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

[8] For single-equation framework, time-specific effect is zero.

[9] Rogoff and Stavrakeva (2008) compare the forecasting power of the monetary model, the Taylor rule model and a structural model based on the Backus-Smith optimal risk sharing condition.

where p_{0t} is the log of the U.S. price level, and p_{it} is the log of the price level of country *i*. I use the real-time CPI to measure of the national price level. Substituting PPP fundamentals (2) into equation (1), I use the resultant equation for forecasting.

3.2 Taylor Rule Fundamentals

When central banks set the interest rate according to the Taylor rule, the linkage between the exchange rate and a set of fundamentals can be examined. According to Taylor (1993), central banks set the monetary policy as:

$$i_t^* = \pi_t + \phi(\pi_t - \pi^*) + \gamma y_t^g + r^* \tag{3}$$

where i_t^* is the target for the short-term nominal interest rate, π_t is the inflation rate, $\pi^*$is the target level of inflation, y_t^g is the output gap, or percent deviation of actual output from an estimate of its potential level, and $r^*$is the equilibrium level of the real interest rate. It is assumed that the target for the short-term nominal interest rate is achieved within the period, so that there is no distinction between the actual and target nominal interest rate.

The parameters π^* and r^* in equation (3) can be combined into one constant term $\mu = r^* - \phi\pi^*$ and we have:

$$i_t^* = \mu + \lambda\pi_t + \gamma y_t^g \tag{4}$$

where $\lambda = 1 + \phi$. If the central bank sets the target the level of exchange rate to make PPP hold, equation (4) becomes:

$$i_t^* = \mu + \lambda\pi_t + \gamma y_t^g + \delta q_t \tag{5}$$

where q_t is the real exchange rate. The central bank increases (decreases) the nominal interest rates if the exchange rate depreciates (appreciates) from its equilibrium value under PPP assumption in the Taylor rule. Allowing the interest rate to achieve its target level within the period:

$$i_t = \mu + \lambda\pi_t + \gamma y_t^g + \delta q_t \tag{6}$$

and i_t is the nominal interest rate. Subtracting the Taylor rule equation for the foreign country from that for the base country, the U.S. (denoted by "0"), equation (6) becomes:

$$i_{0t} - i_{it} = \lambda(\pi_{0t} - \pi_{it}) + \gamma\left(y_{0t}^g - y_{it}^g\right) + \delta(s_{it} + p_{it} - p_{0t}) \tag{7}$$

Imposing the uncovered interest rate parity condition $E_t s_{it+1} = i_{it} - i_{0t} + s_{it}$, the expected change in nominal exchange rates is equal to the interest differential:

$$E_t s_{it+1} = \lambda(\pi_{0t} - \pi_{it}) + \gamma\left(y_{0t}^g - y_{it}^g\right) + \delta(s_{it} + p_{it} - p_{0t}) + s_{it} \tag{8}$$

Molodtsova and Papell (2009) refer to the specification (8) as homogenous asymmetric Taylor rule with no smoothing. They estimate the parameters $\lambda, \gamma,$ and δ in equation (8) country-by-country in a rolling regression framework. Rather than estimating the coefficients, I follow the approach developed by Engel, Mark, and West (2008), who posit a Taylor rule such that λ=1.5, γ=0.1 and δ=0.1. Imposing fixed coefficients for all the countries is preferable for two reasons. First, increasing the number of parameters to be estimated in a panel may reduce the efficiency of forecasts and bring noise to the system. Second, this approach provides a better comparison of forecasts obtained with real-time data and those obtained with ex-post revised data in Engel, Mark, and West (2008). The Taylor rule fundamentals to be used in forecasting equation (1) become:

$$f_{it} = 1.5(\pi_{0t} - \pi_{it}) + 0.1\left(y_{0t}^g - y_{it}^g\right) + 0.1(s_{it} + p_{it} - p_{0t}) + s_{it} \tag{9}$$

It is well known in the literature that the uncovered interest rate parity condition does not hold in the short run. With an error correction specification, the exchange rate forecasting model, $E_t s_{it+k} - s_{it} = \beta_k (f_{it} - s_{it})$, is used to generate out-of-sample forecasts both at the short-horizon (where k=1) and the long-horizon (where k=16).

4. Out-of-Sample Forecasting

4.1 Estimation

To produce out-of-sample forecasts, the sample has to be split into two components, in-sample and out-of-sample. The in-sample component is updated recursively to estimate the parameters in equation (1) within both single-equation and panel frameworks. For single equation estimation, the parameters (constant and β) are estimated country-by-country with OLS. For panel estimation, the parameters (country-specific effects, time specific effects, and β) are estimated by least squares dummy variable (LSDV) method.

Following Mark and Sul (2001) and Engel, Mark, and West (2008), the predictive regression is estimated through 1982:Q4. For k=1 (k=16), the predictive regression is used to forecast 1-step-ahead (16-step-ahead) exchange rate returns in 1983:Q1 (1986:Q4). Then, the in-sample component is updated recursively by extending the sample up to 1983:Q1 and equation (1) is re-estimated at each step. For k=1 (k=16), the predictive regression is used to forecast 1-step-ahead (16-step-ahead) exchange rate returns in 1983:Q2 (1987:Q1), and the loop continues until the last observation. At the end, 105 forecasts for k=1 and 90 overlapping forecasts for k=16 are derived with both PPP and Taylor rule fundamentals.

One crucial point for multi-period ahead forecasts in the panel framework is that the time effect needs to be forecasted. For k-period ahead forecasts, the time effect in period t+k is calculated by taking the recursive mean of the time effect until period t, such as $\hat{\theta}_{t+k} = \frac{1}{t}\sum_{j=1}^{t}\hat{\theta}_j$.

4.2 Comparisons of Forecasts Based on MSPE

To compare the out-of-sample forecasting ability of the two nested models, this study focuses on the minimum mean-squared prediction error (MSPE) approach, which became dominant in the literature after Meese and Rogoff (1983a, 1983b). Forecasts of linear and random walk models are calculated as:

Linear Model: $\Delta\hat{s}_{it+k} = \hat{\zeta}_i + \left(\frac{1}{t}\sum_{j=1}^{t}\hat{\theta}_j\right) + \hat{\beta}z_{it}$

Driftless Random Walk: $\Delta\hat{s}_{it+k} = 0$ (10)

Random Walk with Drift: $\Delta\hat{s}_{it+k} = \hat{\alpha}_{it}$

where $\hat{\alpha}$ is the estimated drift term.[10] Taking the difference between actual and predicted values of exchange rates gives the prediction error. The MSPE approach selects a model which has significantly smaller MSPE than the random walk with or without the drift.

4.3 Out-of-Sample Test Statistics

To measure the relative forecast accuracy of the linear model against the driftless random walk and the random walk with drift, I use two alternative test statistics: the Diebold-Mariano and West (DMW) and the Clark-West (CW) statistics.

4.3.1 The Diebold-Mariano and West (DMW) Test

Suppose that a martingale difference process and a linear model are given as:

---

[10] The recursive mean of the time effect in parenthesis for the linear model is removed in the single-equation case.

Model 1: $y_t = e_t$

Model 2: $y_t = X_t^{'}\beta + e_t$ where $E_{t-1}(e_t) = 0$

where the dependent variable is the change in the exchange rate. Under the null hypothesis, population parameter $\beta = 0$ and exchange rate follows a random walk. For simplicity let us concentrate on one-step-ahead forecasting. Assume that sample size is *T*+1; the first *R* observations are used for estimation and *P* is equal to the number of forecasts. So we have, *T*+1=*R*+*P,* where *T*+1=145, *R*=40 and *P*=105 for one-step-ahead forecasting. Information prior to t is used to forecast for period *t*=*R*, *R*+1, *R*+2, …, *T*. The first forecast is for the period *R*+1 and the final forecast is for the period *T*+1.

The estimated forecasts for the random walk and the structural model are 0 and $X_{t+1}^{'}\hat{\beta}_t$ and $\hat{\beta}_t$ is the regression estimate of β_t. After estimating the forecasts, the respective prediction errors for the models are $\hat{e}_{1,t+1} = y_{t+1}$ and $\hat{e}_{2,t+1} = y_{t+1} - X_{t+1}\hat{\beta}_t$. Thus, the sample MSPE's of the two models become:

$$\hat{\sigma}_1^2 = P^{-1}\sum_{t=T-P+1}^{T} y_{t+1}^2 \quad \text{and} \quad \hat{\sigma}_2^2 = P^{-1}\sum_{t=T-P+1}^{T}(y_{t+1} - X'_{t+1}\hat{\beta}_t)^2 \tag{11}$$

Diebold and Mariano (1995) and West (1996) construct a t-type statistics which is assumed to be asymptotically normal and the population MSPE's are equal under the null. Defining the following equations,

$$\hat{f}_t = \hat{e}_{1,t}^2 - \hat{e}_{2,t}^2$$

$$\bar{f} = P^{-1}\sum_{t=T-P+1}^{T}\hat{f}_{t+1} = \hat{\sigma}_1^2 - \hat{\sigma}_2^2 \tag{12}$$

$$\hat{V} = P^{-1} \sum_{t=T-P+1}^{T} (\hat{f}_{t+1} - \bar{f})^2$$

The DMW test statistic is

$$DMW = \frac{\bar{f}}{\sqrt{P^{-1}\hat{V}}} \tag{13}$$

The asymptotic DMW test works fine with non-nested models. However, the size properties of the asymptotic DMW test have been widely criticized for nested models. Clark and McCracken (2001, 2005) and McCracken (2007) show that the limiting distribution of the DMW test for nested models under the true null is not standard normal. Undersized DMW tests cause too few rejections of the null and may miss the statistical significance of the linear exchange rate model against the random walk.

4.3.2 The Clark- West (CW) Test

Clark and West (2006, 2007) show that the sample difference between the MSPE's of two nested models in DMW test is biased downward from zero in favor of the random walk.

$$\hat{\sigma}_1^2 - \hat{\sigma}_2^2 = P^{-1} \sum_{t=T-P+1}^{T} \hat{f}_{t+1} = P^{-1} \sum_{t=T-P+1}^{T} y_{t+1}^2 - P^{-1} \sum_{t=T-P+1}^{T} (y_{t+1} - X'_{t+1}\hat{\beta}_t)^2 = 2P^{-1} \sum_{t=T-P+1}^{T} y_{t+1} X'_{t+1}\hat{\beta}_t - P^{-1} \sum_{t=T-P+1}^{T} (X'_{t+1}\hat{\beta}_t)^2 \tag{14}$$

Under the null hypothesis, the exchange rate follows a random walk, such that $e_{1,t+1} = e_{2,t+1} = y_{t+1}$. Since the independent variables are not correlated with the disturbance term, the first term in equation (14) is equal to zero.[11] Clark and West (2006, 2007) show that $-P^{-1} \sum_{t=T-P+1}^{T} (X'_{t+1}\hat{\beta}_t)^2 < 0$ because estimating the parameters of the alternative model under the

[11] $\sum_{t=T-P+1}^{T} y_{t+1} X'_{t+1}\hat{\beta}_t$ is zero, because the equality of $e_{1,t+1} = e_{2,t+1}$ under null hypothesis suggests that $E(y_{t+1}X'_{t+1}\hat{\beta}_t) = E(e_{1,t+1}X'_{t+1}\hat{\beta}_t) = E(e_{2,t+1}X'_{t+1}\hat{\beta}_t)$. Since $E(e_{2,t+1}X'_{t+1}) = 0$ by assumption, we have $E(e_{2,t+1}X'_{t+1})E(\hat{\beta}_t) = E(e_{2,t+1}X'_{t+1}\hat{\beta}_t) = 0$.

true null (which are zero) brings noise into the forecasting process. Clark and West (2006) recommend an adjusted DMW statistic that adjusts for the negative bias in the difference between the two MSPE. Defining the adjustments as follows,

$$\hat{f}_{t+1}^{ADJ} = \hat{e}_{1,t+1}^{2} - \left[\hat{e}_{2,t+1}^{2} - (X'_{t+1}\hat{\beta}_t)^2\right]$$

$$\bar{f}^{ADJ} = P^{-1} \sum_{t=T-P+1}^{T} \hat{f}_{t+1}^{ADJ} = \hat{\sigma}_1^2 - \left[\hat{\sigma}_2^2 - P^{-1} \sum_{t=T-P+1}^{T} (X'_{t+1}\hat{\beta}_t)^2\right] \tag{15}$$

$$\hat{V} = P^{-1} \sum_{t=T-P+1}^{T} (\hat{f}_{t+1}^{ADJ} - \bar{f}^{ADJ})^2$$

the CW test statistic is

$$CW = \frac{\bar{f}^{ADJ}}{\sqrt{P^{-1}\hat{V}^{ADJ}}} \tag{16}$$

The CW test has become one the most popular out-of-sample test statistic in the exchange rate literature. However, Rogoff and Stavrakeva (2008) show that the CW test cannot always be interpreted as a minimum MSPE test as the DMW test. Their study presents a proof that in the presence of forecast bias, the null hypothesis of the CW and the DMW tests are not necessarily the same.[12] If one can reject the null of CW test, the true nature of exchange rate does not follow a random walk. Nevertheless, even if the true model follows some other model rather than a random walk, one can still apply the DMW statistics to test whether the random walk and the structural model have equal MSPEs.

4.4 Bootstrapping Out-of-Sample Test Statistics

Size distortions of the DMW test in small samples can be reduced by bootstrapping the finite sample distribution of the test statistics. Kilian (1999) state that unlike asymptotic critical

[12] In the presence of the scale bias, the null hypothesis of the CW and the DMW tests are different.

values, correctly specified (maintaining the cointegration between the exchange rate and fundamentals under the null hypothesis) bootstrap critical values adapt for the increase in the dispersion of the finite-sample distribution of the test statistic. Kilian (1999) also suggest that the bootstrap is appropriate for multi-period ahead forecasts. Based on simulation evidence, Li and Maddala (1997) and Li (2000) also indicate bootstrapped tests have smaller size distortions and higher test power than asymptotic tests in cointegrating systems. Howbeit, Berkowitz and Kilian (2000) emphasize the importance of bootstrapping type implemented to preserve cointegrating relationships in the data. They argue that cointegration appears to be a parametric notion and parametric bootstraps are more accurate than non-parametric ones.

Mark and Sul (2001) and Rogoff and Stavrakeva (2008) apply bootstrapped out-of-sample tests to detect forecasting ability of linear exchange rate models against random walk in a panel framework. The bootstrap methods are similar in both studies. Mark and Sul (2001) implement parametric bootstrap and estimate error correction equations with seemingly unrelated regressions (SURs); however, Rogoff and Stavrakeva (2008) use semi-parametric bootstrap and estimate error correction equations with country specific OLS regressions.

Having insignificant bootstrapped DMW test statistics in certain cases, as opposed to highly significant asymptotic CW test, Rogoff and Stavrakeva (2008) criticize the asymptotic CW test to be oversized and has less power than the bootstrapped DMW test in the presence of forecast bias.[13] Oversized asymptotic CW test would cause too many rejections of the null hypothesis that exchange rate does not follow a random walk. It may detect spurious statistical significance and favor the alternative, structural exchange rate model. In this paper, I evaluate the out-of-sample predictive ability of exchange rate fundamentals based on bootstrapped critical values for CW and DMW tests.

[13] In the technical appendix of Clark and West (2007), the unadjusted power of the bootstrapped DMW test is higher than that of the asymptotic CW test for recursive regressions with one-step-ahead forecasts.

The particular bootstrap method used (which imposes cointegration restriction between the exchange rate and the fundamentals) used in this study is as follows:

$$\Delta s_{it} = \varepsilon_{it} \tag{17}$$

$$\Delta z_{it} = \alpha_i + t + \gamma_i z_{it-1} + \sum_{j=1}^{d} \delta_{ij} \Delta s_{it-j} + \sum_{j=1}^{l} \zeta_{ij} \Delta z_{it-j} + u_{it}$$

where s_{it} is the nominal exchange rate and z_{it} is the deviation of exchange rate from fundamental as defined in equation (1). $\Delta s_{it} = s_{it} - s_{it-k}$ and $\Delta z_{it} = z_{it} - z_{it-k}$ where k is the forecast horizon, α is a constant and t is a trend. To control for autocorrelation in the error correction equation (ECE) lags of Δs_{it} and Δz_{it} are included. Akaike's information criterion is used for each country to determine the optimum number of d and l and to figure out whether to include a constant or a trend or both in the ECE. The sum of coefficients on lags of Δz_{it} is restricted to 1. After specifying the data generating process for each country, equation (17) is estimated with country specific OLS regressions.

The next step to perform the bootstrap is to configure the residual vectors. First, each residual vector at time t is constructed as $\hat{v}_t = (\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t}, ..., \hat{\varepsilon}_{Nt}, \hat{u}_{1t}, \hat{u}_{2t}, ..., \hat{u}_{Nt})'$. Second, $(2N \times 1)$ residual vectors $\hat{v}_1, \hat{v}_2, ..., \hat{v}_T$ are split into non-overlapping blocks of 4 adjacent observations. Let $\hat{w}_1 = (\hat{v}_1, \hat{v}_2, \hat{v}_3, \hat{v}_4)$ be the first block, and L be the number of blocks. I resample the residual blocks $\{\hat{w}_1, \hat{w}_2, ..., \hat{w}_L\}$ with replacement and simulate bootstrap observations of s_{it} and z_{it} recursively, where sample averages are chosen as starting values.

This residual resampling scheme has two important features that need to be addressed. First, constructing the residual vectors as $\hat{v}_t = (\hat{\varepsilon}_t, \hat{u}_t)'$ account for the cross-sectional dependence in the estimated residuals across countries. Second, resampling in blocks maintain the serial correlation properties of the residuals.

To reduce the bias caused by the initial values of the recursion, the first 100 observations are thrown away and a new sample is created. Applying the estimation procedure again, test statistics are calculated with the pseudo-data. This process is repeated 1000 times and semi-parametric bootstrap distribution is derived. Since the tests considered are one-sided tests, the p-values of DMW and CW tests are the percentage of the bootstrapped distribution above the estimated test statistic using the realized data.

5. Empirical Results

This section compares one- and 16-quarter-ahead out-of-sample performance of the linear exchange rate model with PPP and Taylor rule fundamentals to that of the random walk model with and without drift using a newly constructed real-time dataset. The tables report the MSPE ratio, the ratio of the MSPE of the structural model to that of the random walk, and the DMW and CW test statistics with their respective bootstrapped p-values. A significant DMW or CW test statistic implies that the linear exchange rate model outperforms the random walk with or without the drift out-of-sample.

5.1 PPP Fundamentals

One-quarter-ahead single-equation forecasting results with the PPP model are presented in Table 2. No evidence of out-of-sample predictability is found with the PPP model against the driftless random walk for any exchange rate. The out-of-sample performance of the PPP model improves against the random walk with drift. Short-term predictability is found for Canada with the

CW test and for 5 countries (Canada, Germany, Japan, Netherlands, and the U.K.) with the DMW test at the one-quarter horizon.

Panel one-quarter-ahead forecasts using PPP fundamentals in Table 3 are slightly better compared to single-equation forecasts in Table 2. The exchange rate model with PPP fundamentals using panel data significantly outperforms the driftless random walk only for Japan. The evidence of predictability of the PPP model with panel estimation, just like in the single-equation case, increases against the random walk with drift at one-quarter horizon. Short-term predictability is found for 4 out of 9 countries (Australia, Canada, Japan, and Sweden) with the CW test and for Australia and Sweden with the DMW test.

The low predictive power of the PPP model at the one-quarter horizon using panel and single-equation estimations is not surprising. Existing studies concerning the half-life of PPP, the expected number of years for a PPP deviation to decay by 50%, find half-lives of around 2.5 years.[14] Accounting for the slow adjustment of real exchange rates in advanced economies, one would expect the predictive ability of PPP model to be low at short horizons.

Sixteen-quarter-ahead out-of-sample forecasts with the PPP model and single-equation estimation are presented in Table 4. The evidence of long-term predictability is stronger compared to one-quarter-ahead forecasts using the single-equation framework with rejections of the random walk null found for 4 countries (France, Germany, Netherlands, and Sweden) with the CW test. More evidence of long-term predictability is found against the random walk with drift. Out-of-sample exchange rate predictability is found for 7 out of 9 countries (Australia, Canada, France, Germany, Japan, Netherlands, and Sweden) with the CW test and for 3 countries (Australia, Canada and Netherlands) with the DMW test. The out-of-sample predictability of the PPP model with a

[14] See Wu (1996), Papell (1997, 2002), Murray and Papell (2002), Choi, Mark and Sul (2006) for details concerning the half-lives of PPP deviations.

single-equation framework is clearly improved at the l6-quarter horizon compared to one-quarter horizon.

The PPP model performs best with the panel specification at the 16-quarter horizon. As reported in Table 5, the evidence of predictability is found for 7 out of 9 countries (Australia, Canada, France, Germany, Japan, Netherlands, and Sweden) with the CW test, and for 5 out of 9 countries (Australia, Canada, Germany, Japan, and Netherlands) with the DMW test against the driftless random walk. Panel forecasts at long horizon against the random walk with drift are, in fact, striking. Out-of-sample predictability is found for all the countries in the sample regardless of which test statistic is used. Because the persistence of deviation from PPP across countries is relatively homogenous, panel estimation becomes more efficient and the predictability of the panel exchange rate model with PPP fundamentals is much higher than the single-equation framework.

5.2 Taylor Rule Fundamentals

Following Engel, Mark, and West (2008), predictive regressions using Taylor rule model are estimated where the coefficients on inflation, output gap, and real exchange rate are fixed at certain values. One-quarter-ahead single-equation forecasts with Taylor rule are reported in Table 6. Evidence of short-term predictability is found only for Japan against the driftless random walk. The exchange rate model with Taylor fundamentals works much better against the random walk with drift. Evidence of out-of-sample predictability found for 3 out of 9 countries (Australia, Japan, and Sweden) with the CW test and for 5 out of 9 countries (Australia, Canada, Japan, Netherlands, and Sweden) with the DMW test.

Comparing Tables 6 and 7, the performance of Taylor rules does not get improved by panel estimation. Presence of substantial heterogeneity causes time-series regression forecasts to be superior as suggested in Mark and Sul (2012).[15] One-quarter ahead forecasting results for the Taylor rule model with a panel framework are reported in Table 7. No evidence of out-of-sample predictability is found against the driftless random walk regardless of which test statistic is used. The results are stronger against the random walk with drift. Evidence of predictability is found for 3 out of 9 countries (Australia, Japan, and Sweden) with the CW test and for 5 out of 9 countries (Australia, Germany, Japan, Netherlands, and Sweden) with the DMW test.

Table 8 presents 16-quarter-ahead single-equation forecasts using the Taylor rule model. Evidence of long-term predictability is found only for Germany with CW test against the driftless random walk. The single equation forecasts with the Taylor rule model perform better against the random walk with drift. Evidence of long-term predictability is found for Sweden with the CW test and for 4 out 9 countries (Australia, Japan, Netherlands, and Sweden) with the DMW test.

Panel forecasts with the Taylor rule model at the 16-quarter horizon perform poorly. As reported in Table 9, no evidence of either long-term predictability is found against the random walk, with or without drift, for any of the countries in the sample regardless of which test statistic is used. Low forecasting power of the Taylor rule model at the long horizon is reasonable because central banks target short-term nominal interest rates. These results are in accord with previous work using revised or quasi-real-time data. Molodtsova and Papell (2009) report that the evidence of short term predictability disappears at longer horizons with a single equation Taylor rule model, and Engel, Mark, and West (2008) do not find more evidence of predictability with panel models.

6. Conclusions

15 Mark and Sul (2012) show that pooling does not dominate time-series regression in out-of-sample forecasting when the heterogeneity is great.

The purpose of this paper is to investigate how real-time data affects out-of-sample predictability of PPP and Taylor rule exchange rate models at short and long horizons using single-equation and panel frameworks. The vast majority of empirical studies on exchange rate forecasting over the post-Bretton Woods period use ex-post revised data, which contain future information that was not available to policymakers and market participants at the time the forecasts were made. Therefore, it cannot be used to evaluate predictability of exchange rate models out-of-sample. Forecasts with real-time data, however, do not contain any unrealized future information in the information set of the linear model, mimic the information set of market agents as closely as possible, and thus can be used to construct a true out-of-sample forecast.

Engel, Mark, and West (2008) find that panel error-correction exchange rate models with PPP fundamentals are able to produce large improvements in out-of-sample forecasting at longer horizons. Because they use ex-post revised data, the exchange rate models in their study contain future information that was not available to market participants. The results in this paper show that panel estimation increases the predictability of the PPP model relative to single-equation estimation. Having relatively homogenous deviations from PPP across countries cause panel estimation to be more efficient and estimating the predictive regression with panel data increases the forecasting power of the PPP model. At the 16-quarter horizon, evidence of predictability is found with panel estimation for 7 out of 9 countries with the CW test and 5 out of 9 countries with the DMW test against the driftless random walk and for all of the countries against the random walk with drift regardless of which test statistic is used. One-quarter-ahead forecasts of the exchange rate model with PPP fundamentals are weaker than long-horizon forecasts. Strong predictability of the PPP model at longer-horizons with panel estimation is in accord with estimated half-lives of PPP deviations of around 2.5 years, and confirms the findings in Engel, Mark, and West (2008).

Molodtsova and Papell (2009), using ex-post revised data to calculate inflation and quasi-real-time data to estimate output gaps, find evidence of out-of-sample exchange rate predictability with the Taylor rule model at short horizon using single-equation estimation. While quasi-real-time data does not contain future observations, it captures revisions which are not available to market participants in real-time. Therefore, quasi-real time data also cannot be used to produce true out-of-sample forecasts. Out-of-sample forecasting exercises in our study show that the predictability of the Taylor rule model is higher at the short horizon than at the long horizon as in Molodtsova and Papell (2009). Evidence of short-term predictability with the single-equation Taylor rule model is found for 1 out of 9 countries with both test statistics against the driftless random walk, and for 3 out of 9 countries with the CW test and 5 out of 9 countries with the DMW test against the random walk with drift. Since, central banks target short-term nominal interest rates, low predictive ability of Taylor rules at the long-horizon is not surprising. In contrast to PPP model, panel Taylor rule exchange rate models are unable to improve the forecasts compared to single-equation estimation, which is consistent with the results in Engel, Mark, and West (2008). Weak panel performance of Taylor rules suggest existence of substantial heterogeneity in adjustment to equilibrium.

References

Berkowitz, J., Kilian, L., 2000. Recent Developments in Bootstrapping Time Series. Econometric Reviews 19, 1-48.

Cerra, V., Saxena, S. C., 2010. The Monetary Model Strikes Back: Evidence from the World. Journal of International Economics 81, 184-196.

Cheung, Y.-W., Chinn, M. D., Pascual, A.G., 2005. Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive? Journal of International Money and Finance 24, 1150-1175.

Choi, C.-Y., Mark, N., Sul, D., 2006. Unbiased Estimation of the Half-Life to PPP Convergence in Panel Data. Journal of Money, Credit, and Banking 38, 921-938.

Clarida, R., Gali, J., Gertler, M., 1998. Monetary Rules in Practice: Some International Evidence. European Economic Review 42, 1033-1067.

Clark, T. E., McCracken, M. W., 2001. Tests of Equal Forecast Accuracy and Encompassing for Nested Models. Journal of Econometrics 105, 671-110.

Clark, T. E., McCracken, M. W., 2005. Evaluating Direct Multi-Step Forecasts. Econometric Reviews 24, 369-404.

Clark, T. E., West, K. D., 2006. Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis. Journal of Econometrics 135, 155-186.

Clark, T. E., West, K. D., 2007. Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. Journal of Econometrics 138, 291-311.

Diebold, F., Mariano, R., 1995. Comparing Predictive Accuracy. Journal of Business and Economic Statistics 13, 253-263.

Dornbusch, R., 1976. Expectations and Exchange Rate Dynamics. Journal of Political Economy 84, 1161-1176.

Engel, C., Mark, N. C., West, K. D., 2008. Exchange Rate Models Are Not as Bad as You Think. In: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), NBER Macroeconomics Annual 2007. University of Chicago Press, pp. 381-441.

Faust, J., Rogers, J. H., Wright, J. H., 2003. Exchange Rate Forecasting: the Errors We've Really Made. Journal of International Economics 60, 35-59.

Frankel, J. A., 1979. On The Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials. American Economic Review 69, 601-622.

Gerdesmeier, D., Mongelli, F. P., Roffia, B., 2007. The Eurosystem, the US Federal Reserve and the Bank of Japan: Similarities and Differences. Journal of Money, Credit, and Banking 39, 1785-1819.

Gourinchas, P.-O., Rey, H., 2007. International Financial Adjustment. Journal of Political Economy 115, 665-703.

Hodrick, R. J., Prescott, E.C., 1997. Postwar U.S. Business Cycles: An Empirical Investigation. Journal of Money, Credit, and Banking 29, 1-16.

Ince, O., Papell, D. H., 2013. The (Un)Reliability of Real-Time Output Gap Estimates with Revised Data. Economic Modelling 33, 713-721.

Kilian, L., 1999. Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions? Journal of Applied Econometrics 14, 491-510.

Li, H., 2000. The Power of Bootstrap Based Tests for Parameters in Cointegrating Regressions. Statistical Papers 41, 197-210.

Li, H., Maddala, G., 1997. Bootstrapping Cointegrating Regressions. Journal of Econometrics 80, 297-318.

Mark, N. C., 1995. Exchange Rate and Fundamentals: Evidence on Long-Horizon Predictability. American Economic Review 85, 201-218.

Mark, N. C., Sul, D., 2001. Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel. Journal of International Economics 53, 29-52.

Mark, N. C., Sul, D., 2012. When Are Pooled Panel-Data Regression Forecasts of Exchange Rates More Accurate than the Time-Series Regression Forecasts? In: James, J., Marsh, I.W., Sarno, L. (Eds.), Handbook of Exchange Rates. John Wiley and Sons Inc., pp. 256-281.

McCracken, M. W., 2007. Asymptotics for Out-of-Sample Tests of Granger Causality. Journal of Economerics 140, 719-752.

Meese, R. A., Rogoff, K., 1983a. Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample? Journal of International Economics 14, 3-24.

Meese, R. A., Rogoff, K., 1983b. The Out of Sample Failure of Empirical Exchange Rate Models. In: Frenkel, J.A. (Ed.), Exchange Rates and International Macroeconomics. University of Chicago Press, pp. 6-105.

Murray, C. J., Papell, D. H., 2002. The Purchasing Power Parity Persistence Paradigm. Journal of International Economics 56, 1-19.

Molodtsova, T., Papell, D. H., 2009. Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals. Journal of International Economics 77, 167-180.

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

Molodtsova, T., Nikolsko-Rzhevskyy, A., Papell, D. H., 2011. Taylor Rules and the Euro. Journal of Money, Credit, and Banking 43, 535-552.

Papell, D. H., 1997. Searching for Stationarity: Purchasing Power Parity Under the Current Float. Journal of International Economics 43, 313-332.

Papell, D. H., 2002. The Great Appreciation, the Great Depreciation, and the Purchasing Power Parity Hypothesis. Journal of International Economics 57, 51-82.

Rogoff, K., Stavrakeva, V., 2008. The Continuing Puzzle of Short Horizon Exchange Rate Forecasting. National Bureau of Economic Research Working Paper 14071.

Rossi, B., 2005. Testing Long-Horizon Predictive Ability with High Persistence, and the Meese-Rogoff Puzzle. International Economic Review 46, 61-92.

Rossi, B., 2006. Are Exchange Rates Really Random Walks? Some Evidence Robust to Parameter Instability. Macroeconomic Dynamics 10, 20-38.

Taylor, J. B., 1993. Discretion versus Policy Rules in Practice. Carnegie-Rochester Conference Series on Public Policy 39, 195-214.

Watson, M., 2007. How Accurate Are Real-Time Estimates of Output Trends and Gaps? Federal Reserve Bank of Richmond Economic Quarterly 93, 143-161.

West, K. D., 1996. Asymptotic Inference about Predictive Ability. Econometrica 64, 1067-1084.

Wu, Y., 1996. Are Real Exchange Rates Nonstationary? Evidence from a Panel-Data Test. Journal of Money, Credit, and Banking 28, 54-63.

Table 1. Descriptive Statistics

A. INFLATION

	Real-Time Data				Revised Data			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Australia	5.901	4.005	-1.348	16.188	5.888	3.974	-0.389	16.188
Canada	4.541	3.236	0.000	11.908	4.605	3.235	0.000	11.986
France	4.862	3.980	0.345	14.004	4.854	3.990	0.193	14.084
Germany	2.884	1.813	-1.078	7.470	2.862	1.823	-1.002	7.152
Italy	7.478	5.608	1.293	22.048	7.444	5.628	1.384	22.011
Japan	2.909	4.410	-1.396	20.403	2.910	4.522	-1.396	22.505
Netherlands	3.439	2.539	-1.227	10.318	3.406	2.558	-1.206	10.312
Sweden	5.163	3.879	-0.810	13.768	5.149	3.910	-0.995	13.694
U.K.	6.479	5.076	1.034	22.530	6.403	5.088	1.034	23.433
U.S.	4.525	2.843	1.207	13.504	4.544	2.853	1.207	13.543

B. OUTPUT GAP

	Real-Time Data				Revised Data			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Australia	-0.707	2.563	-9.209	8.894	0.063	2.919	-9.719	7.040
Canada	-0.899	2.045	-6.275	2.943	0.109	3.405	-13.339	6.152
France	-1.437	2.080	-9.639	3.150	0.065	2.413	-9.064	6.476
Germany	-1.025	1.928	-6.616	2.359	0.056	2.775	-8.078	6.344
Italy	-2.077	2.335	-8.289	2.671	0.057	3.266	-12.332	9.768
Japan	-1.526	2.764	-17.783	2.864	0.103	4.017	-12.105	11.792
Netherlands	-1.369	1.535	-7.142	2.862	-0.020	2.300	-7.179	5.419
Sweden	-0.558	2.177	-5.576	6.672	0.167	3.392	-11.496	6.692
U.K.	-0.949	1.825	-5.972	3.362	0.089	2.330	-5.914	7.654
U.S.	-0.909	2.577	-9.393	4.279	0.199	3.365	-11.478	6.970

Notes: The statistics reported for each variable are: Mean, the mean, SD, the standard deviation, Min, and Max, the minimum and maximum values. The data is for 1973:Q1-2009:Q1.

Table 2. Single Equation 1-Quarter-Ahead Forecasts Using PPP Fundamentals

	No Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	1.0219	0.1770	0.5910	-0.6003	0.4730
Canada	1.0058	0.2108	0.4240	-0.3231	0.2380
France	1.0223	0.1191	0.3990	-0.6880	0.3120
Germany	0.9975	0.7763	0.2290	0.0987	0.1290
Italy	1.0739	0.2634	0.6870	-1.5571	0.7980
Japan	0.9918	0.9504	0.3530	0.3329	0.1700
Netherlands	1.0012	0.5748	0.3110	-0.0486	0.1710
Sweden	1.0217	0.6944	0.3700	-0.8252	0.5080
U.K.	1.0179	0.2829	0.4100	-0.5022	0.3050
	Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	1.0068	0.3595	0.2850	-0.2165	0.1960
Canada	0.9878	**1.1801**	**0.0800**	**0.6040**	**0.0360**
France	0.9978	0.6150	0.2130	0.1194	0.1230
Germany	0.9897	1.0355	0.1070	**0.4836**	**0.0500**
Italy	1.0197	-1.1403	0.8340	-1.4620	0.7430
Japan	0.9964	0.6673	0.1870	**0.2881**	**0.0980**
Netherlands	0.9925	0.9005	0.1340	**0.3739**	**0.0710**
Sweden	0.9991	1.1124	0.1410	0.0473	0.1670
U.K.	0.9970	0.5721	0.1970	**0.1643**	**0.0890**

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.

Table 3. Panel 1-Quarter-Ahead Forecasts Using PPP Fundamentals

	No Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.9870	1.0847	0.3030	0.4723	0.1830
Canada	0.9938	0.8499	0.2640	0.2249	0.1880
France	1.0084	0.0201	0.3770	-0.4093	0.2990
Germany	0.9876	1.1311	0.1680	0.4429	0.1320
Italy	1.0519	0.3837	0.5510	-0.9149	0.7320
Japan	0.9726	**2.0353**	**0.0870**	**1.5247**	**0.0400**
Netherlands	0.9926	0.8988	0.1880	0.3024	0.1370
Sweden	0.9907	0.7144	0.3360	0.3540	0.1870
U.K.	1.0042	0.6631	0.3040	-0.1231	0.3070
	Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.9725	**1.6333**	**0.0710**	**1.0345**	**0.0960**
Canada	0.9761	**1.5325**	**0.0850**	0.6836	0.1730
France	0.9842	1.1510	0.1360	0.6060	0.2510
Germany	0.9799	1.3525	0.1070	0.7937	0.1620
Italy	0.9987	0.5847	0.2830	0.0490	0.4330
Japan	0.9772	**1.4911**	**0.0810**	0.8698	0.1330
Netherlands	0.9840	1.1972	0.1370	0.6509	0.1910
Sweden	0.9687	**1.8830**	**0.0500**	**1.0867**	**0.0930**
U.K.	0.9836	1.2211	0.1180	0.7057	0.2280

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.

Table 4. Single Equation 16-Quarter-Ahead Forecasts Using PPP Fundamentals

			No Drift		
	MSPE ratio	CW	P-value	DMW	P-value
Australia	1.3393	1.1002	0.2680	-0.7412	0.5240
Canada	1.0883	0.6955	0.2550	-0.2140	0.2400
France	2.0458	**2.0382**	**0.0180**	-0.9436	0.4540
Germany	1.4950	**2.6105**	**0.0080**	-0.4662	0.3230
Italy	2.9815	0.6797	0.4050	-0.9610	0.5280
Japan	1.0300	1.0890	0.3020	-0.0642	0.3300
Netherlands	0.8441	**2.4793**	**0.0180**	0.1788	0.1180
Sweden	2.9874	**2.2454**	**0.0220**	-1.5296	0.6960
U.K.	2.1290	0.5006	0.3480	-0.6647	0.3610
			Drift		
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.7913	**4.3300**	**0.0000**	**0.7716**	**0.0170**
Canada	0.7596	**2.2683**	**0.0150**	**0.8363**	**0.0250**
France	1.1539	**3.8812**	**0.0020**	-0.2461	0.2980
Germany	1.3551	**2.7336**	**0.0050**	-0.3689	0.5770
Italy	1.1597	1.0209	0.1500	-0.1992	0.2510
Japan	1.0042	**2.5483**	**0.0030**	-0.0093	0.1820
Netherlands	0.7356	**2.8317**	**0.0040**	**0.3482**	**0.0730**
Sweden	1.7764	**4.2205**	**0.0010**	-1.0049	0.8000
U.K.	1.0219	0.9871	0.1930	-0.0269	0.2160

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.

Table 5. Panel 16-Quarter-Ahead Forecasts Using PPP Fundamentals

	No Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.7614	**1.6545**	**0.0920**	**0.6718**	**0.0990**
Canada	0.6711	**1.6328**	**0.0540**	**0.8875**	**0.0390**
France	0.8818	**1.2697**	**0.0560**	0.2028	0.1310
Germany	0.4211	**2.3768**	**0.0170**	**0.9245**	**0.0480**
Italy	2.0116	0.4540	0.4410	-0.7276	0.6040
Japan	0.6032	**2.6244**	**0.0380**	**1.5011**	**0.0430**
Netherlands	0.4651	**2.3258**	**0.0120**	**0.9939**	**0.0290**
Sweden	0.4975	**1.6649**	**0.0910**	0.6658	0.1040
U.K.	1.4518	0.3979	0.3340	-0.3942	0.3810
	Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.4499	**5.6980**	**0.0000**	**2.6217**	**0.0010**
Canada	0.4684	**3.6361**	**0.0000**	**2.0548**	**0.0020**
France	0.4973	**3.3588**	**0.0000**	**1.5287**	**0.0070**
Germany	0.3817	**2.5197**	**0.0050**	**1.0893**	**0.0210**
Italy	0.7825	**1.0278**	**0.0360**	**0.4021**	**0.0980**
Japan	0.5881	**4.1808**	**0.0000**	**1.5983**	**0.0030**
Netherlands	0.4053	**2.7490**	**0.0010**	**1.2681**	**0.0070**
Sweden	0.2958	**3.4463**	**0.0020**	**1.5690**	**0.0100**
U.K.	0.6968	**1.2608**	**0.0310**	**0.5513**	**0.0910**

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.

Table 6. Single Equation 1-Quarter-Ahead Forecasts Using Taylor Rule Fundamentals

	MSPE ratio	CW	P-value	DMW	P-value
No Drift					
Australia	0.9914	1.2537	0.2770	0.2250	0.1870
Canada	1.0060	0.5849	0.4260	-0.2338	0.2950
France	1.0527	-1.3383	0.8850	-1.9355	0.8420
Germany	1.0147	0.6132	0.2610	-0.3617	0.1930
Italy	1.0597	-0.3097	0.7610	-1.4349	0.7930
Japan	0.9652	**1.9937**	**0.0980**	**1.4085**	**0.0160**
Netherlands	1.0048	0.4043	0.3110	-0.1809	0.1300
Sweden	1.0100	1.0844	0.3590	-0.2601	0.3690
U.K.	1.0333	-1.1452	0.9140	-1.6155	0.7970
Drift					
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.9786	**1.6592**	**0.0910**	**1.2498**	**0.0420**
Canada	0.9858	1.1145	0.1970	**0.8788**	**0.0500**
France	1.0213	-1.5723	0.9180	-1.7368	0.8370
Germany	1.0048	0.6663	0.2600	-0.1424	0.1710
Italy	1.0060	0.7238	0.2620	-0.1340	0.1970
Japan	0.9697	**1.9859**	**0.0190**	**1.6030**	**0.0040**
Netherlands	0.9934	0.7929	0.1550	**0.2549**	**0.0640**
Sweden	0.9773	**1.6497**	**0.0950**	**1.1215**	**0.0430**
U.K.	1.0092	-0.5116	0.6950	-0.7419	0.4270

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.

Table 7. Panel 1-Quarter-Ahead Forecasts Using Taylor Rule Fundamentals

			No Drift		
	MSPE ratio	CW	P-value	DMW	P-value
Australia	1.0014	0.8280	0.4240	-0.0426	0.4200
Canada	1.0094	0.4920	0.4970	-0.4324	0.5090
France	1.0314	-0.5309	0.6310	-1.1586	0.6250
Germany	1.0022	0.3006	0.4100	-0.1204	0.3180
Italy	1.0725	0.1114	0.6670	-1.3378	0.8510
Japan	0.9848	1.2875	0.3070	0.5006	0.2240
Netherlands	1.0056	-0.0354	0.5230	-0.3726	0.3760
Sweden	1.0185	0.3264	0.6200	-0.6734	0.7060
U.K.	1.0292	-0.1793	0.5930	-0.9453	0.5560
			Drift		
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.9884	**2.0224**	**0.0920**	**1.8584**	**0.0780**
Canada	0.9891	1.4148	0.2490	1.2825	0.1590
France	1.0006	0.0641	0.5590	-0.0969	0.4500
Germany	0.9924	1.5397	0.1470	**1.4379**	**0.0830**
Italy	1.0183	-1.3881	0.9350	-1.6899	0.9180
Japan	0.9894	**1.8692**	**0.0330**	**1.7603**	**0.0260**
Netherlands	0.9940	1.1742	0.1510	**1.0690**	**0.0960**
Sweden	0.9856	**2.1221**	**0.0730**	**2.0224**	**0.0490**
U.K.	1.0052	-0.4187	0.6400	-0.6499	0.5510

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.

Table 8. Single Equation 16-Quarter-Ahead Forecasts Using Taylor Rule Fundamentals

	No Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	1.3755	0.3071	0.4580	-0.6126	0.4530
Canada	1.3806	-0.4763	0.6630	-1.3708	0.6860
France	3.1752	0.2036	0.4070	-1.4136	0.5550
Germany	1.0439	**1.5246**	**0.0780**	-0.0514	0.1650
Italy	3.7153	0.0230	0.5870	-1.4449	0.6680
Japan	0.9203	0.7005	0.4560	0.2044	0.2510
Netherlands	1.0795	0.8599	0.2060	-0.1441	0.1700
Sweden	1.5230	0.5261	0.3940	-0.6842	0.4550
U.K.	2.9587	-0.5476	0.6550	-1.3239	0.4770
	Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.8639	0.7102	0.1490	**0.3535**	**0.0940**
Canada	0.9302	0.4608	0.2500	0.3733	0.1020
France	1.5964	1.0529	0.2090	-0.7708	0.4570
Germany	0.8746	1.5925	0.1180	0.1755	0.1490
Italy	1.4247	-0.0450	0.6050	-0.5894	0.5280
Japan	0.8974	0.6737	0.1440	**0.2698**	**0.0910**
Netherlands	0.8547	1.3971	0.1050	**0.3329**	**0.0940**
Sweden	0.7617	**1.4365**	**0.0560**	**0.6234**	**0.0560**
U.K.	1.3117	-0.3061	0.6110	-0.4753	0.3860

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.

Table 9. Panel 16-Quarter-Ahead Forecasts Using Taylor Rule Fundamentals

	No Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	1.4467	-0.0038	0.5630	-0.7712	0.5620
Canada	1.4097	-0.6368	0.7500	-1.4443	0.7430
France	2.0068	-0.8928	0.7910	-1.4713	0.7590
Germany	1.0488	0.2465	0.4600	-0.1870	0.3850
Italy	2.5953	-0.2666	0.7200	-1.3622	0.7940
Japan	0.9283	0.7547	0.4440	0.1388	0.3430
Netherlands	1.1738	-0.5536	0.7120	-0.8821	0.5660
Sweden	1.7350	-0.3142	0.6690	-1.0647	0.6920
U.K.	1.8665	-0.4944	0.6240	-0.9460	0.4790
	Drift				
	MSPE ratio	CW	P-value	DMW	P-value
Australia	0.9086	0.3606	0.2650	0.2513	0.1550
Canada	0.9498	0.3397	0.3420	0.2625	0.1880
France	1.0088	0.1881	0.5950	-0.0258	0.5070
Germany	0.8786	0.7962	0.3330	0.5564	0.2400
Italy	0.9953	0.1283	0.6430	0.0105	0.4970
Japan	0.9052	0.2868	0.2570	0.1881	0.2020
Netherlands	0.9293	0.7348	0.2550	0.4532	0.1800
Sweden	0.8678	0.5384	0.2030	0.3829	0.1550
U.K.	0.8275	0.5368	0.2380	0.4247	0.1360

Notes: The table reports the MSPE ratio, defined as the ratio of MSPEs of the linear exchange rate model to that of the benchmark model (random walk with and without the drift), the CW statistics and the DMW statistics for the tests of equal MSPEs. All reported tests are one-sided. Bold font denotes the p-value of respective test statistic significant at 10 % level based on semi-parametric bootstrap. Starting in 1973:Q1, I estimate recursive regressions with a 40-quarter initial window to predict exchange rate changes from 1983:Q1 to 2009:Q1.