EFWW
7/10/07
11:00 AM

Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals

Tanya Molodtsova[*] and David H. Papell[**]

University of Houston

January 23, 2007

Abstract

An extensive literature that studied out-of-sample performance of empirical exchange rate models following Meese and Rogoff's (1983a) seminal paper has not yet convincingly overturned their result of no out-of-sample predictability of exchange rates. The recent empirical research by Cheung, Chinn and Pascual (2005) concludes that none of the standard models of exchange rate determination consistently outperforms the random walk model at any horizon. This paper re-evaluates the short and long-term predictability of empirical exchange rate models using Clark and West's (2005a) recently developed inference procedure for testing for equal predictability of two nested models. We extend the conventional set of models of exchange rate determination – the flexible price monetary model and its two building blocks (uncovered interest rate parity and purchasing power parity) – by investigating predictability of two models that incorporate a Taylor-rule interest rate reaction function. The out-of-sample performance of the models is assessed at 1 to 36 month horizons for a set of 12 currencies over the post-Bretton Woods float. The paper provides evidence of short-term exchange rate predictability, which could not be observed using the standard inference procedure prevailing in the previous literature. The evidence of predictability is much stronger with Taylor rule models than with conventional models. There is no significant evidence of increased predictability of economic models relative to a random walk at long horizons.

We are grateful to Menzie Chinn, Todd Clark, Lutz Kilian, Mike McCracken, Moto Shintani, Vadym Volosovych, Jian Wang, Ken West, and participants at the 2006 Econometric Society Summer Meetings, International Economics and Finance Society sessions at the 2006 SEA and 2007 ASSA meetings, and Texas Camp Econometrics XI for helpful comments and discussions.

[*] Department of Economics, University of Houston, Houston, TX 77204-5019. Tel: +1 (832) 858-2076 Email: tvmolodt@mail.uh.edu

[**] Department of Economics, University of Houston, Houston, TX 77204-5882. Tel/Fax: +1 (713) 743-3807/3798. Email: dpapell@uh.edu

1. Introduction

The failure of open-economy macro theory to explain exchange rate behavior using economic fundamentals has prevailed in the international economics literature since the seminal papers by Meese and Rogoff (1983a, 1983b), who examine the out-of-sample performance of three empirical exchange rate models during the post-Bretton Woods period. Using monthly data from March 1973 through November 1980 for generating one-to-twelve month horizon predictions, they find that the random walk model produces consistently more accurate forecasts than the empirical exchange rate models on the basis of the root mean squared error (RMSE) comparison. The authors conclude that economic models of exchange rate determination of the 1970's vintage do not perform better than a naïve "no change" model.

Starting with Mark (1995), a number of studies have found evidence of greater predictability of economic exchange rate models at longer horizons. Mark considers a model where the changes in the log spot exchange rate depend on the deviation of the current log exchange rate from its "fundamental value", which is motivated by the flexible-price monetary model of exchange rate determination and is determined as a linear combination of log relative money stocks and real incomes in the two countries. Chen and Mark (1996) assess the out-of-sample performance of the three alternative fundamentals proposed in the literature: those implied by the purchasing power parity (PPP), the uncovered interest rate parity (UIRP) and the flexible-price monetary model. Both studies analyze four currencies using quarterly data from 1973 to the beginning of the 1990s and find the evidence of greater predictability of economic models at long horizons. The authors conclude that monetary fundamentals have the most predictive power. Chinn and Meese (1995), using an error correction version of the model, also come to a conclusion of long-term exchange rate predictability.

The findings that macroeconomic fundamentals have predictive power relative to the random walk and out-of sample performance of economic models increases with the forecast horizon have been questioned in subsequent research by Kilian (1999), Faust, Rogers and Wright (2003), and Berkowitz and Giorgianni (2001).[1] The recent comprehensive study by Cheung, Chinn and Pascual (2005) examines the out-of-sample performance of the interest rate parity, monetary, productivity-based and behavioral exchange rate models and concludes that none of the models consistently outperforms the random walk at any horizon.

The relevant literature on exchange rate predictability compares out-of-sample performance of two models (linear fundamental-based model and a random walk) on the basis of different measures. The most commonly used measure of predictive ability is mean squared prediction error (MSPE). To evaluate out-of-sample performance of the models based on the MSPE comparison, the tests for equal predictability of two

[1] Some of the authors point out the problems with the bootstrap procedure employed by Mark (1995), while the others argue that the results are not robust to even modest changes in the sample period.

non-nested models, introduced by Diebold and Mariano (1995) and West (1996), is most often used (henceforth, DMW tests).[2]

The first contribution of this paper is to apply a new inference procedure for testing the null of equal predictive ability of a linear econometric model and a martingale difference model proposed by Clark and West (2005a, 2005b), which we call the CW procedure. This methodology is preferable to the standard DMW procedure when the two models are nested. The test statistic appropriately takes into account that under the null the sample MSPE of the alternative model is expected to be greater than that of the random walk model. The simulation evidence in Clark and McCracken (2001, 2005b), McCracken (2006) and Corradi and Swanson (2005) demonstrates that when comparing MSPE's of two nested models mechanical application of the procedures in Diebold and Mariano (1995) and West (1996) leads to non-normal test statistics and the use of standard normal critical values usually results in very poorly sized tests, with far too few rejections of the null.[3] In addition to being severely undersized, the standard DMW procedure demonstrates very low power, which makes this statistic ill-suited for detecting departures from the null[4]. Rossi (2005) documents the existence of size distortions of the DMW tests by revisiting the Meese and Rogoff puzzle. While the paper suggests a possible way to solve this problem by adjusting critical value of the tests, the resulting statistic has low power.

McCracken (2004) develops an asymptotic approximation that rationalizes the non-normal distribution of the commonly used test statistic revealed by those simulations. Previous literature has implicitly or explicitly assumed that since the difference in population MSPE's is zero, the difference in sample MSPE should be approximately zero. However, Clark and West (2005a) demonstrate analytically that the sample MSPE from the alternative model is expected to be greater than that of the null. The proposed statistic adjusts for the upward shift in the sample MSPE of the alterative model. The simulations in Clark and West (2005a) suggest that the inference made using their asymptotically normal critical values results in properly-sized tests.[5]

An alternative for investigating exchange rate predictability, following Mark (1995) and Kilian (1999), is to construct critical values using simulation or bootstrap techniques. While Clark and West (2005a) find that the gains from bootstrapping in terms of size and power are negligible compared to using adjusted CW statistics with normal critical values, we do not choose one technique over the other. We compare inference

---

[2] Under the null of equal predictive accuracy, the DMW statistic is assumed to be zero and has an asymptotic standard normal distribution. One of the most highly cited studies that use the DMW statistic for testing for equal predictive ability of the models is Mark (1995).

[3] McCracken (2004) shows that using standard normal critical values for the DMW statistic results in severely undersized tests, with tests of nominal 0.10 size generally having actual size less than 0.02.

[4] According to simulations in Clark and West (2005a), the probability of detecting the failure of the null using the CW procedure is substantial (around 0.6 as opposed to 0.09 with DMW and standard normal critical values) but not overwhelming.

[5] West (2005) provides a summary of recent literature on asymptotic inference about forecasting ability.

using the CW methodology to that using the DMW statistic with both standard normal and bootstrapped critical values. While the standard DMW statistic is undersized, the bootstrapped DMW statistic is a proper comparison to the CW statistic because the bootstrap accounts for the non-standard asymptotic distribution implied by the nested model comparisons. The resulting test tends to have accurate size.[6]

Why are we concerned with undersized tests? In many cases, undersized tests are much less of a problem than oversized tests. In the case of exchange rate predictability, however, the typical result is that the random walk null cannot be rejected in favor of the model-based alternative. Using undersized tests, such as the unadjusted DMW statistic with standard normal critical values, could lead to the incorrect conclusion that the random walk forecasts better than the economic model when the model has statistically significant predictive power.

We examine the role that the inference methodology has played in drawing conclusions about exchange rate predictability. To our knowledge, three previous papers address this issue. Clark and West (2005a) examine the predictability of the UIRP fundamental-based model for 4 countries vis-à-vis the U.S. dollar: Japan, Switzerland, Canada and the U.K. Using the adjusted statistic they find that the economic model significantly outpredicts the random walk for Canada and Switzerland at a 1-month horizon. Using the CW inference procedure, Gourinchas and Rey (2005) find that the ratio of net exports to net foreign assets forecasts movements in FDI-weighted and trade-weighted exchange rates better than the no-change model at 1 to 16 quarter horizon. Alquist and Chinn (2006) examine out-of-sample performance of the sticky-price monetary model, UIRP model and a measure of external imbalances suggested by Gourinchas and Rey. The CW procedure rejects the null of no predictability for UIRP at long horizons and a transformation of net exports variable performs well at short horizons.

We use the adjusted test statistic to evaluate the out-of-sample performance of exchange rate models based on three different fundamentals (purchasing power parity, uncovered interest rate parity, and monetary) for 12 currencies vis-à-vis the U.S. dollar in the post-Bretton Woods period. We find statistically significant evidence of exchange rate predictability at the 1-month horizon for 6 of the 12 currencies using at least one of the models. The predictive power of all of the models decreases sharply with the forecast horizon, and disappears entirely after the 6-month horizon.

The second contribution of this paper is to extend the set of conventional macroeconomic exchange rate models to models with Taylor rule fundamentals. In Taylor's (1993) original formulation, the rule posits that the Fed sets the nominal interest rate based on the current inflation rate, the inflation gap - the difference between inflation and the target inflation rate, the output gap - the difference between GDP and potential GDP, and the equilibrium real interest rate. Assuming that the foreign central bank follows a similar rule and

[6] Kilian (1999) provides simulation evidence for the size and power of the bootstrapped DMW statistic. According to simulations in Clark and West (2005a), the CW statistic with standard normal inference has slightly smaller size and about the same power as the DMW statistic with bootstrapped critical values.

that UIRP holds, we construct a *symmetric* model with Taylor rule fundamentals where the exchange rate is determined by relative expected inflation gaps and relative output gaps. Following Clarida, Gali, and Gertler (1998), we can also posit that the foreign central bank includes the difference between the exchange rate and the target exchange rate, defined by PPP, in its Taylor rule and construct an *asymmetric* model with Taylor rule fundamentals where the exchange rate is also determined by its deviation from PPP.

Although the role of interest rate differentials for exchange rate determination has been previously studied, the Taylor rule approach to interest rate modeling is a relatively unexplored area. It introduces a multivariate structure into exchange rate behavior, which generates a richer set of dynamics and has the potential for producing interest rate forecasts with higher predictive ability. Mark (2005) considers Taylor rule interest rate reaction functions for Germany and the U.S. and estimates the real dollar-mark exchange rate path assuming that the exchange rate is priced by uncovered interest rate parity. He provides evidence that the interest rate differential can be modeled as a Taylor rule differential and the real dollar-mark exchange rate is linked to the Taylor rule fundamentals, which may provide a resolution for the exchange rate disconnect puzzle. Engel and West (2006) constructs a "model-based" real exchange rate as the present value of the difference between home and foreign output gaps and inflation rates, and find a positive correlation between the "model-based" rate and the actual dollar-mark real exchange rate.

We evaluate the out-of-sample performance of models with symmetric and asymmetric Taylor rule fundamentals using the Clark and West adjustment of the DMW statistic. In order to construct Taylor rule fundamentals, we need to define the output gap, and we use deviations from a linear trend, deviations from a quadratic trend, and the Hodrick-Prescott filter. In accord with recent work on estimating Taylor rules for the United States, we define potential GDP using "semi-real time" trends which are updated each period.

The results provide more evidence of short-run exchange rate predictability than we found by using monetary, PPP, and UIRP fundamentals and support the idea that the Taylor rule plays an important role in explaining exchange rate behavior. For the symmetric model, we find statistically significant evidence of exchange rate predictability at the 1-month horizon for 6 of the 12 currencies using at least one of the output gap measures. For the asymmetric model, we find similar evidence of exchange rate predictability for 7 of the 12 currencies. As with the earlier models, the predictive power of the models with Taylor rule fundamentals decreases sharply with the forecast horizon.

The finding of no out-of-sample exchange rate predictability at long horizons is consistent with the results of Kilian (1999) and Wang (2006). Kilian (1999) provides simulation evidence of no increased long-horizon predictability in a conventional VEC setting. Using conventional test with bootstrapped critical values, he finds that the predictability of linear monetary model decreases or stays the same with the horizon. Wang (2006) connects the Engel and West (2005) explanation for low predictability to the study of long-horizon exchange rate regressions. He finds that, under the degree of persistence of the change of fundamentals observed in the data, the exchange rate behaves more like a random walk at long horizons than at short horizons, making it harder to detect statistical evidence in favor of economic models over the random walk in long-horizon regressions than in short-horizon regressions. This explanation is consistent with our findings; using a variety of models, we find strong evidence of predictability at the 1-month horizon, some evidence of predictability at the 6-month horizon, but no predictability at longer horizons.

2. Exchange Rate Models

A large number of exchange rate models have been used to evaluate the link between exchange rates and macroeconomic fundamentals over the post-Bretton Woods period. We focus on the most widely used and easily implementable form of the model, which represents a change in (the logarithm of) the nominal exchange rate as a function of its deviation from its fundamental value. Thus, the h-period-ahead change in the log exchange rate can be modeled as a function of its current deviation from its fundamental value.

(1) $$s_{t+h} - s_t = a_h + \beta_h z_t + \nu_{t+h,t},$$

where $$z_t = f_t - s_t$$

The variable s_t is the log of the nominal exchange rate determined as the domestic price of foreign currency and f_t is the long-run equilibrium level of the nominal exchange rate determined by macroeconomic fundamentals. This approach to exchange rate modeling has been used by Mark (1995) and other subsequent research.

We evaluate the role of inference methodology in studying exchange rate predictability. In line with previous research, we use the random walk as a benchmark model for comparison. We evaluate the out-of-sample performance of the model (1) at 1- to 36-month horizons. The remainder of this section describes five different formulations of the fundamentals. We start by looking at the fundamentals implied by models that are based on Taylor rules for monetary policy. We look at symmetric Taylor rule fundamentals where central banks in both countries adjust the interest rate in response to the deviations in expected inflation and output gap and asymmetric Taylor rule fundamentals where the foreign country also targets deviations of its exchange rate from the desired level, which is determined by PPP. We also consider three conventional fundamentals which have been used in previous research on exchange rate predictability. These are long-run values of the exchange rate implied by the flexible-price monetary model and its two building blocks – PPP and UIRP.

2.1 Taylor Rule Fundamentals

We examine the linkage between the exchange rates and a set of fundamentals that arise when central banks set the interest rate according to the Taylor rule. Assuming that UIRP holds, the exchange rate is determined by the expected inflation differential, the output gap differential, and, if one or both of the central banks target the exchange rate, the deviation from PPP.

Following Taylor (1993), the monetary policy rule postulated to be followed by central banks is

(2) $$i_t = \pi_t + \delta(\pi_t - \pi^T) + \alpha y_t^{gap} + r$$

where i_t is the short-term nominal interest rate, π_t is the inflation rate, π^T is the target level of inflation, y_t^{gap} is the output gap, or percentage difference between actual and potential output at time t, and r is the equilibrium level of the real interest rate.

According to the Taylor rule (2), the central bank raises the nominal interest rate if inflation rises above its desired level and/or output is above potential output. The target level of output deviation from its long-run trend y_t^{gap} is 0 because, according to the natural rate hypothesis, output cannot permanently exceed potential output. The target level of inflation is positive because it is generally believed that deflation is much worse for an economy than low inflation.

The parameters in π^T and r in equation (2) can be combined into one constant term, which leads to the following equation,

(3) $$i_t = \mu + (1+\delta)\pi_t + \alpha y_t^{gap}$$

We consider two types of Taylor rules. The first specification assumes that both the home and foreign countries determine their interest rate according to the same Taylor rule of the standard form (3). We refer to this model as the symmetric Taylor rule model. Assuming that the coefficients in the Taylor rule, as well as the target inflation rates and equilibrium real interest rates, are equal in the two countries, a similar expression can be obtained for the foreign country,

(4) $$i_t^* = \mu + (1+\delta)\pi_t^* + \alpha y_t^{*gap}$$

Subtracting the interest rate reaction function for the foreign country (4) from that for the home country (3), the following equation for the interest rate differential can be obtained:

(5) $$i_t - i_t^* = \alpha(y_t^{gap} - y_t^{*gap}) + (1+\delta)(\pi_t - \pi_t^*)$$

where asterisks denote foreign country variables.

The second Taylor-rule specification assumes that the interest rate reaction function for the foreign country explicitly includes the exchange rate:

(6) $$i_t^* = \mu - \gamma(s_t - \bar{s}_t^*) + (1+\delta)\pi_t^* + \alpha y_t^{*gap}$$

where $\bar{s}_t^*$ is the target level of the exchange rate in the foreign country and $0 < \gamma < 1$. To simplify the model, we assume that the monetary authorities set the target level of the exchange rate to make PPP hold:

(7) $$\bar{s}_t^* = (p_t - p_t^*)$$

Since we defined s_t as the dollar price of one unit of foreign currency, the rule (6) means that the foreign country increases its interest rate when its currency depreciates relative to the target level of the exchange rate. Similar monetary policy rules for different countries are estimated in Clarida, Gali and Gertler (1998). Equation (3) for the home country does not include the exchange rate because there is no evidence that the U.S. has adopted an exchange rate target.[7]

Subtracting (6) from (3) we obtain,

(8) $$i_t - i_t^* = \gamma(s_t - \bar{s}_t^*) + \alpha(y_t^{gap} - y_t^{*gap}) + (1+\delta)(\pi_t - \pi_t^*)$$

Substituting the target level of exchange rate (7) into equation (8) leads to,

(9) $$i_t - i_t^* = \gamma(s_t - (p_t - p_t^*)) + \alpha(y_t^{gap} - y_t^{*gap}) + (1+\delta)(\pi_t - \pi_t^*)$$

Assuming that the expected rate of depreciation equals to the interest rate differential, or that uncovered interest rate parity (UIRP) holds,

(10) $$E(\Delta s_{t+h}) = (i_t - i_t^*)$$

This implies the following fundamental value of the log exchange rate in equation (1),

(11) $$f_t = (i_t - i_t^*) + s_t$$

Finally, we substitute the interest rate differentials in equations (5) and (9) into the expression for UIRP fundamental (11) and estimate equation (1) for each of the two Taylor rule models in rolling regressions.

2.2 Monetary Fundamentals

We select the flexible-price monetary model as representative of 1970's vintage models. The monetary approach determines the exchange rate as a relative price of the two currencies, and models exchange rate behavior in terms of relative demand for and supply of money in the two countries. The long-run money market equilibrium in the domestic and foreign country is given by:

(12) $$m_t = p_t + ky_t - \lambda i_t$$

(13) $$m_t^* = p_t^* + k^* y_t^* - \lambda^* i_t^*,$$

[7] As discussed by Engel and West (2005), this specification would still be valid if the U.S. had an exchange rate target in its interest rate reaction function.

where m_t, $p_{t,}$ and y_t are the logs of money supply, price level and income and i_t is the level of interest rate in period t; asterisks denote foreign country variables.

The monetary model assumes purchasing power parity, according to which exchange rates in the two countries will move to balance the prices:

(14) $$s_t = p_t - p_t^*,$$

where s_t is the log of nominal exchange rate determined as the domestic price of foreign currency.

Subtracting equation (13) from equation (12), using the PPP condition (14) to solve for the exchange rate, and assuming that countries are homogenous in terms of income elasticities and interest rate semi-elasticities of money supplies, we obtain the following equation,

(15) $$s_t = (m_t - m_t^*) - k(y_t - y_t^*) + \lambda(i_t - i_t^*)$$

According to equation (15), an increase in the money supply differential between the domestic and foreign countries leads to a depreciation of the domestic currency. On the other hand, an increase in relative income of the domestic country creates additional demand for domestic money. In response to this, domestic residents cut their consumption forcing prices to fall, which leads through PPP to an appreciation of the domestic currency.

Assuming that UIRP holds and substituting the one-period-ahead interest rate differential in (14), the following expression for the exchange rate is obtained,

(16) $$s_t = (m_t - m_t^*) - k(y_t - y_t^*) + \lambda E(\Delta s_{t+1})$$

Iterating forward equation (16) and using the identity $E(\Delta s_{t+1}) = E(s_{t+1}) - s_t$ leads to,

(17) $$s_t = \frac{1}{1+\lambda}\sum_{i=1}^{\infty}\left(\frac{\lambda}{1+\lambda}\right)^t E_t\left[(m_{t+i} - m_{t+i}^*) - k(y_{t+i} - y_{t+i}^*)\right]$$

However, expression (17) represents the solution to (16) only in the absence of rational bubbles. In general, equation (16) has multiple solutions that satisfy

(18) $$s_t^0 = s_t + \nu_t\ ,$$

where the rational bubble term satisfies

(19) $$E_t[\nu_{t+1}] = \left(\frac{1+\lambda}{\lambda}\right)\nu_t$$

We take the approach used in Mark (1995) and other subsequent research and denote the fundamental value of the exchange rate as

(20) $$f_t = (m_t - m_t^*) - k(y_t - y_t^*)$$

We construct the monetary fundamental with a fixed value of income elasticity, k, equal to 0, 0.5, and 1. In this paper, we report only the results for k=0, because the pattern of the results is robust to changes in k, but statistical significance is weaker for values of k equal to 0.5 and 1[8]. We substitute the monetary fundamentals (20) into (1), and forecast exchange rates using rolling regressions.

2.3 Purchasing Power Parity Fundamentals

As a basis of comparison, we examine the predictive power of PPP fundamentals. There has been extensive research on PPP in the last decade, and a growing body of literature finds that long-run PPP holds in the post-1973 period[9]. Since the monetary model is build upon PPP but assumes additional restrictions, comparing the out-of-sample performance of the two models is a logical exercise. Mark and Sul (2001) use panel-based forecasts and find evidence that the linkage between exchange rates and monetary fundamentals is tighter than that between exchange rates and PPP fundamentals.

Under PPP fundamentals,

(21) $$f_t = (p_t - p_t^*)$$

where p_t is the log of the national price level. We use the CPI as a measure of national price levels. We substitute the PPP fundamentals (21) into (1), and forecast exchange rates using rolling regressions.

2.4 Uncovered Interest Rate Parity Fundamentals

The second building block of the monetary model is UIRP. Recent evidence indicates that exchange rate movements are consistent with UIRP in the long-run, but not in the short-run.[10] Under the UIRP condition (15), the expected *h*-period ahead change in log exchange rate is equal to the nominal interest rate differential. This determines UIRP fundamentals

(22) $$f_t = (i_t - i_t^*) + s_t$$

We use 3-month Treasury bill rate as a measure of the short-term interest rate, substitute the UIRP fundamentals (22) into (1), and forecast exchange rates using rolling regressions. Instead of using the exact formulation of the expected change in the exchange rate implied by UIRP as in Alquist and Chinn (2006), we use the interest rate differentials as fundamentals as in Clark and West (2005a).

---

[8] It is noted by Mark (1995) that as long as the results are robust to the variations in *k*, movements in the money supply differential are relatively more important that movements in the income differential.

[9] See Papell (2006) for a recent example.

[10] See Chinn and Meredith (2004) and Chinn (2006).

3. Forecast Comparison

3.1 Data

The models are estimated using monthly data from March 1973, the beginning of the floating exchange rate period, through December, 1998 for European Monetary Union countries and October, 2004 for the rest of the countries[11]. The currencies we consider are the Japanese yen, Swiss franc, Australian dollar, Canadian dollar, British pound, Swedish kronor, Danish kroner, Deutsche mark, French franc, Italian lira, Dutch guilder, and Portuguese escudo. The exchange rates defined as the US dollar price of a unit of foreign currency are taken from the Federal Reserve Bank of Saint Louis database. [12]

The primary source of data used to construct macroeconomic fundamentals is the IMF's *International Financial Statistics* (IFS) database[13]. We use M1 to measure the money supply for most of the countries. We use M0 for the U.K. and M2 for Italy and Netherlands, because M1 data is unavailable for these countries. Using M2 as a measure of the money supply provides similar results. We use the seasonally adjusted industrial production index (IFS line 66) as a proxy for countries' national income since GDP data are available only at the quarterly frequency. [14] The price level in the economy is measured by consumer price index (IFS line 64). The inflation rate is the annual inflation rate, measured as the 12-month difference of the CPI. Our choice of countries reflects our intention to examine exchange rate behavior for major industrialized economies with flexible exchange rates over the sample.

The output gap depends on the measure of potential output. Since there is no presumption about which definition of potential output is used by central banks in their interest rate reaction functions, we consider percentage deviations of actual output from a linear time trend, a quadratic time trend, and a Hodrick-Prescott (1997) (HP) trend as alternative definitions. In order to mimic as closely as possible the information available to the central banks at the time the decisions were made, we use semi-real time data in the output gap estimation. For a given period t, we use only the data points up to t-1 to construct the trend. Thus, in each period the OLS regression is re-estimated adding one additional observation to the sample.[15]

---

[11] Some of the models are estimated using shorter spans of data because of data unavailability. The footnotes for the tables list these exceptions.

[12] The UIRP model is estimated using data from January 1980 for Japan and Switzerland, March 1977 for Italy, and July 1975 for Germany because the interest rate data are not available prior to those periods. Also the Treasury bill series stops in October 2003 for Japan, March 2004 for Sweden, and June 2004 for Australia. The monetary model is estimated using the data from December 1977 for France, December 1974 for Italy, and December 1979 for Portugal because of the lack of money supply data prior to those periods.

[13] The complete Data Appendix and data files are available at the author's web-site: www.uh.edu/~dpapell

[14] The industrial production series for Australia and Switzerland, and the CPI series for Australia, which were available only quarterly, are transformed into monthly periodicity using the "quadratic-match average" option in Eviews 4.0.

[15] We call this semi-real time data because, while the trend is updated each period, the data incorporate revisions that were not available to the central banks at the time decisions were make. True real time data is not available for most of the countries that we study over the entire floating rate period.

3.2 Estimation and Forecasting

We construct 1- to 36-month ahead forecasts for the linear regression models (1) with each of the five fundamentals described above. We use data over the period March 1973 - February 1982 for estimation and reserve the remaining data for out-of-sample forecasting. Let us concentrate for simplicity on one-step-ahead predictions. To evaluate the out-of-sample performance of the models, we estimate them by OLS in rolling regressions and construct CW and DMW statistics. Each model is initially estimated using the first 120 data points and the one-period-ahead forecast is generated. Then, we drop the first data point, add an additional data point at the end of the sample, and re-estimate the model. A one period-ahead forecast is generated at each step. [16]

3.3 Forecast Comparison Based on MSPE

Each model's out-of-sample forecasting ability is compared to that of the martingale difference process using an adjusted test statistic, which is constructed as described in Clark and West (2005a). We are interested in comparing the mean square prediction errors from the two nested models. The first model is a zero mean martingale difference process, while the other is a linear model.

Model 1: $y_t = \varepsilon_t$

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

Suppose we have a sample of T+1 observations. The last P observations are used for predictions. The first prediction is made for the observation R+1, the next for R+2, …, the final for T+1. We have T+1=R+P, R=120, P=260 for non-EU countries and 190 for EU countries. To generate prediction for period t=R, R+1, …, T, we use the information available prior to t. Let $\hat{\beta}_t$ is a regression estimate of β_t that is obtained using the data prior to t. The one-step ahead prediction for model 1 is 0, and $X_{t+1}\hat{\beta}_t$ for model 2. The sample forecast errors from the models 1 and 2 are $\hat{e}_{1,t+1} = y_{t+1}$ and $\hat{e}_{2,t+1} = y_{t+1} - X_{t+1}\hat{\beta}_t$, respectively. The corresponding MSPE's for the two models are $\hat{\sigma}_1^2 = P^{-1}\sum_{t=T-P+1}^{T} y_{t+1}^2$ and $\hat{\sigma}_2^2 = P^{-1}\sum_{t=T-P+1}^{T}(y_{t+1} - X'_{t+1}\hat{\beta}_t)^2$.

[16] We use out-of-sample rather than in-sample methods for comparison with the extensive literature following Meese and Rogoff (1983a), estimate rolling rather than recursive regressions because the CW statistic is only valid for the former, and choose a rolling window of 120 observations to estimate alternative forecast models following the empirical exercise in Clark and West (2005a). Inoue and Kilian (2004) advocate using in-sample rather than out-of-sample methods and using recursive methods for out-of-sample forecasting.

We are interested in testing the null hypothesis of no predictability against the alternative that exchange rates are linearly predictable.[17] Thus,

$$H_0 : \sigma_1^2 - \sigma_2^2 = 0$$

$$H_1 : \sigma_1^2 - \sigma_2^2 > 0$$

Under the null, the population MSPE's are equal. We need to use the sample estimates of the population MSPE's to draw the inference. The procedure introduced by Diebold and Mariano (1995) and West (1996) uses sample MSPE's to construct a t-type statistics which is assumed to be asymptotically normal. To construct the DMW statistic, let

$$\hat{f}_t = \hat{e}_{1,t}^2 - \hat{e}_{2,t}^2 \qquad \text{and} \qquad \bar{f} = P^{-1} \sum_{t=T-P+1}^{T} \hat{f}_{t+1} = \hat{\sigma}_1^2 - \hat{\sigma}_2^2$$

Then, the DMW test statistics is computed as follows,

$$DMW = \frac{\bar{f}}{\sqrt{P^{-1}\hat{V}}}, \qquad \text{where} \qquad \hat{V} = P^{-1} \sum_{t=T-P+1}^{T} (\hat{f}_{t+1} - \bar{f})^2 \tag{23}$$

Clark and West (2005a) demonstrate analytically that the asymptotic distributions of sample and population difference between the two MSPE's are not identical, namely the sample difference between the two MSPE's is biased downward from zero. This means that using the test statistic (23) with standard normal critical values is not advisable.

It is straightforward to show that the sample difference between the two MSPE's is uncentered under the null.

$$\hat{\sigma}_1^2 - \hat{\sigma}_2^2 = P^{-1} \sum_{t=T-P+1}^{T} \hat{f}_{t+1} = P^{-1} \sum_{t=T-P+1}^{T} y_{t+1}^2 - P^{-1} \sum_{t=T-P+1}^{T} (y_{t+1} - X'_{t+1}\hat{\beta}_t)^2 = 2P^{-1} \sum_{t=T-P+1}^{T} y_{t+1} X'_{t+1}\hat{\beta}_t - P^{-1} \sum_{t=T-P+1}^{T} (X'_{t+1}\hat{\beta}_t)^2 \tag{24}$$

Under the null, the first term in (24) is zero, while the second one is greater than zero by construction. Therefore, under the null we expect the MSPE of the naïve no-change model to be smaller than that of a linear model. The intuition behind this result is the following. If the null is true, estimating the alternative model introduces noise into the forecasting process because it is trying to estimate parameters which are zero in population. In finite samples, use of the noisy estimate of the parameters will lead to higher

[17] We use the term "predictability" as a shorthand for "out-of-sample predictability" in the sense used by Clark and West (2005a,b), rejecting the null of a zero slope in the predictive regression in favor of the alternative of a nonzero slope.

estimated MSPE. As a result, the sample MSPE of the alternative model will be higher by the amount of estimation noise.

To properly adjust for this shift, we construct the corrected test statistic as described in Clark and West (2005a) by adjusting the sample MSPE from the alternative model by the amount of the bias in the second term of equation (24). This adjusted CW test statistic is asymptotically standard normal. When the null is a martingale difference series Clark and West (2005a, 2005b) recommend adjusting the difference between MSPE's as described above and using standard normal critical values for inference[18]. We also calculate the commonly used in the literature DMW statistic, and compare it to standard normal. It is shown by the above algebraic exercise and supported by many simulations in Clark and McCracken (2001, 2005b), McCracken (2004) and Corradi and Swanson (2005) that using standard normal critical values with DMW statistic results in severely undersized tests.

3.4 Inference

Since we are simultaneously testing multiple hypotheses, inference based on p-values is likely to be contaminated. This issue arises because we have 9 different models of 12 bilateral exchange rates using 2 tests yielding 210 test statistics. As a result, it is not surprising to find some evidence of predictability using 5 or 10 percent p-value thresholds. To increase the reliability of our results for each of the tests, we also make inference about each test based on its q-value, a measure of significance of a test statistic which takes into account that we are conducting multiple tests simultaneously. These q-values are constructed using methodology discussed in Storey (2003) and implemented by McCracken and Sapp (2005) for tests of predictive ability. Although the use of q-values is increasingly popular in the statistics literature, this methodology is still relatively unknown in the economics literature. Below we provide a brief description of the methodology and the intuition behind it.[19]

The q-value associates each test with its own measure of significance. While the p-value is a measure of significance in terms of the false positive rate, the q-value is a measure in terms of the false discovery rate. Given a decision rule for rejecting the null, the false positive rate is the rate that true null hypotheses are rejected. The false discovery rate, introduced by Benjamini and Hochberg (1995), is the rate that rejected hypotheses are truly null hypotheses. If a test has a p-value of 5%, we would expect that on average in repeated samples 5% of the true null hypotheses will be rejected. If a statistic has a q-value of 5%, we would expect that on average 5% of the rejected hypotheses are actually true null hypotheses. The q-value threshold can be interpreted as the proportion of rejected hypotheses that turn out to be wrongly rejected.

---

[18] Because the null hypothesis for the CW statistic is a zero mean martingale difference process, we can only test the null that the exchange rate is a random walk, not a random walk with drift. Clark and West (2005a, b) argue that standard normal critical values are approximately correct, event though the statistics are non-normal according to Clark and McCracken (2001), and advocate using them instead of bootstrapped critical values.

[19] This description is largely drawn from McCracken and Sapp (2005).

Consider an experiment in which m hypotheses are being tested. Suppose m_0 and $m_1 = m - m_0$ are the number of cases when the null and alternative hypotheses are true respectively. If we denote the total number of rejections by S, and the total numbers of false and true rejections by F and T respectively, we will obtain the following table.

	Reject	Fail to Reject	Total
Null true	F	$m_0 - F$	m_0
Alternative true	T	$m_1 - T$	m_1
Total	S	$m - S$	m

In a standard situation when only one hypothesis is being tested (m=1), one selects a decision rule which maximizes the power of the test, $E(T/m_1) = Pr(T = 1)$, when $m_1 = 1$, keeping the false positive rate or a probability of a type I error, $E(F/m_0) = Pr(F = 1)$, when $m_0 = 1$, below some threshold α. If the p-value associated with the test statistic is less than or equal to α, we reject the null hypothesis.

When the number of tests is more than one, it is not clear how one should choose the decision rule and, in particular, how to decide which of the m hypotheses correspond to the null or the alternative hypotheses. Applying the standard approach designed for a single test, we can no longer ensure that the false positive rate, $E(F/m_0)$, is less than α but can only ensure that $E(F/m) \leq \alpha$.. This results in too many false rejections of the null. One solution is the Bonferroni correction, which solves the problem by setting the p-value threshold to be α/m instead of α. This ensures that the false positive rate is below the threshold, $E(F/m_0) \leq \alpha$, but does so at the expense of power.

Controlling for the false discovery rate, E(F/S), rather than the false positive rate, $E(F/m_0)$, is useful for multiple hypothesis testing in certain settings. It is especially useful when one is testing many hypotheses and is interested in having a low frequency of false positives among all the rejected hypotheses. Taking into account the nature of our problem, this method is an appealing intermediate alternative to either a liberal single-test approach or a conservative Bonferroni correction. We are more interested in ensuring that our rejections are correct (F/S is small) than in keeping the false positive rate small (F/m_0 is small). Storey (2003) defines the q-value as the minimum possible false discovery rate for which we reject the null, just as the p-value is defined as the minimum possible false positive rate for which we reject the null.

We provide both p- and q-values for each of the 210 test statistics. The primary assumptions we rely on while constructing the q-values are that (1) the p-values are asymptotically valid, (2) the p-values satisfy certain weak dependence conditions, and (3) m is large. The standard normal critical values for the CW statistic are asymptotically valid when we have rolling regressions and a martingale difference null hypothesis. To get asymptotically valid p-values for the DMW statistic, we use the bootstrap. It is less clear that we satisfy the last two assumptions. McCracken and Sapp suggest that smaller sample sizes and stronger dependence imply increasingly conservative (i.e. larger) estimates of the q-values.[20]

4 Empirical Results

4.1 One-Month Ahead Forecasts

To illustrate how the CW statistic is constructed, consider the following exercise. Suppose we estimate the symmetric Taylor rule model with the output gap defined as a deviation from a linear time trend. Let us pick Japan as a representative country for illustration purposes. The unadjusted MSPE of the model is 8.10, which is higher than the MSPE of the random walk of 7.83. Therefore, the DMW statistic is negative, with a t-statistic equal to -0.66. To obtain a test statistic that is centered around zero, we need to adjust the MSPE of the economic model downward. This adjustment term, $P^{-1}\sum_{t=T-P+1}^{T}(X'_{t+1}\hat{\beta}_t)^2$, is equal to 0.52 in our example. After subtracting this adjustment term from the MSPE of the economic model, we obtain the adjusted MSPE which is smaller that that of the random walk and equal to 7.58. Thus, the CW statistic is positive with a t-statistic equal to 1.45. This means that the test based on adjusted difference rejects the null of no exchange rate predictability with the Taylor rule model at 10% significance level.

In tables 1-3 we report the p- and q-values for the DMW and CW statistics for each model and currency. For DMW statistic, we report p-values calculated both based on standard normal critical values (in columns DMW standard), and based on bootstrapped critical values (in columns DMW bootstrap). The p-values for CW statistic are based on standard normal critical values. To construct the q-values, we use asymptotically valid p-values.

Before discussing individual tables it is worthwhile to make a few general observations about the p- and q-values and how they can be used to interpret the results. Using the notation introduced earlier, of the m=210 tests we have 43 cases where the p-values are less than 10% ($S_{10\%}$ = 43), 29 cases where the p-values are less than 5% ($S_{5\%}$ = 29), and 10 cases where the p-values are less than 1% ($S_{1\%}$ = 10). Similarly, of the 210 tests, 10 have q-values less than 10%, and 0 have q-values less than 5 or 1%. Each of the 43, 29 and 10 statistics with p-values less than 10%, 5%, and 1% respectively have q-values less than 27.6%, 19.0%, and 9.8% respectively. Using a 10%, 5%, or 1% p-value threshold implies that at most 27.6%, 19.0%, or 9.8% of the rejections are false, respectively. Thus, when we use the 10%, 5%, or 1% threshold for the p-values, the corresponding q-values indicate that we expect at most $E(F_{10\%}/S_{10\%})$ = 27.6%, $E(F_{5\%}/S_{5\%})$ = 19.0%, or $E(F_{1\%}/S_{1\%})$ = 9.8% of the rejections correspond to false discoveries. The expected number of false positives

[20] The algorithm for constructing q-values for each test and their properties are described in Storey (2002) and Storey and Tibshirani (2003). We use the publicly available software package QVALUE to construct the asymptotically valid q-values associated with each of the 210 tests. A detailed discussion of the software can be found at http://faculty.washington.edu/~jstorey/qvalue/.

$E(F_{10\%}) = 27.6\% \cdot 43 \approx 12$, $E(F_{5\%}) = 19.0\% \cdot 29 \approx 6$, $E(F_{1\%}) = 9.8\% \cdot 10 \approx 1$ is much smaller than the number of rejections implied by 10%, 5%, and 1% rejection thresholds. Since the expected number of false discoveries is reasonably small as a proportion of those rejections, we use 10%, 5%, and 1% p-value rejection thresholds as decision rules for detecting predictive ability in the models despite the existence of multiple testing. Alternatively, we could set a q-value threshold of, say, 10% and reject the null only when the probability of making wrong rejection is less than 10%.

Taylor Rule Fundamentals

Table 1 summarizes the results for 1-month-ahead forecasts of exchange rates using symmetric Taylor rule fundamentals with linear, quadratic and HP trends to estimate potential output. Using the Clark and West procedure, the model with Taylor rule fundamentals significantly outperforms the random walk for 2 out of 12 countries with a linear trend (Canada at 1% and Japan at 10% significance level), for 2 out of 12 countries with a quadratic trend (Canada at 1% and Italy at 5% significance level), and for 6 out of 12 countries with an HP trend (Canada at 1%, U.K., Italy and Portugal at 5%, and France and Japan at 10% level). The model significantly outperforms the random walk with at least one of the output gap specifications for 6 out of 12 currencies. The results also illustrate the well known difficulty of outpredicting the naïve no-change model based on MSPE comparisons. The DMW procedure with standard normal critical values concludes that the random walk performs significantly better than the symmetric Taylor rule model for all of the currencies and output gap specifications. The p-values for DMW statistic calculated using bootstrap suggest more evidence of predictive ability than standard normal p-values, although not as much evidence as with the CW statistic. According to the p-values, the model significantly outperforms the random walk for Canada (at 1% level) with a linear and a quadratic trend, and for 3 out of 12 countries with an HP trend (Canada at 1%, U.K at 5%, and France at 10% level). The model significantly outperforms the random walk with at least one of the output gap specifications for 3 out of 12 currencies.

Short-term predictability increases when we use the Taylor rule where the foreign country targets the exchange rate. Table 2 presents the results for the Taylor rule model with exchange rate targeting. Using the Clark and West procedure, the model significantly outperforms the random walk for 4 out of 12 countries with a linear trend (Denmark and Portugal at 5%, and Canada and Sweden at 10% significance level), for 5 out of 12 countries with a quadratic trend (Canada, Denmark and Portugal at 5%, and Italy and Sweden at 10% significance level) and for 5 out of 12 countries with an HP trend (U.K. and Portugal at 1%, Canada and Italy at 5%, and France at 10% level). The model significantly outperforms the random walk with at least one of the output gap specifications for 7 of the 12 of the countries. Combining the symmetric and asymmetric Taylor rule models, evidence of short-term predictability is found for 8 out of 12 countries: the exceptions being Switzerland, Germany, Australia, and Netherlands. In accord with the results in Table 1, the model cannot significantly outperform the random walk for any output gap specification using the DMW statistic with standard normal inference. The DMW statistic with bootstrapped critical values indicates weaker predictive ability than the CW statistic. According to the p-values, the model significantly outperforms the random walk for 2 out of 12 countries with a linear trend (Canada at 5% and Denmark at 10%) and an HP trend (Canada and U.K. at 5%), and only for Canada (at 5% level) with a quadratic trend. The model significantly outperforms the random walk with at least one of the output gap specifications for 3 out of 12 currencies.

The findings confirm the importance of inference methodology for exchange rate predictability. The CW statistic provides strong evidence of exchange rate predictability based on Taylor rule fundamentals, especially when the asymmetric model with exchange rate targeting is used, while the commonly used DMW statistic does not provide support for the Taylor rule models over the random walk model. This is not a case of two tests providing conflicting results. Since the DMW statistic is severely undersized, and the CW statistic is approximately correctly sized, the results with the CW statistic provide support for exchange rate predictability that would be masked if the (undersized) DMW tests were used. Using the DMW statistic with bootstrapped critical values provides more evidence of exchange rate predictability than the DMW statistic with standard normal critical values, although not as much evidence as with the CW statistic.

Monetary Fundamentals

Table 3 contains the results for 1-month-ahead forecasts of the exchange rates using the monetary, UIRP, and PPP fundamentals described in Section 2. The CW statistic provides some evidence of exchange rate predictability. The monetary model has higher predictive ability for 3 out of 12 countries (Switzerland and Japan at 5% level, and Denmark at 10% level). As with the Taylor rule models, no evidence of predictability is found with the DMW statistic. The monetary model does not produce forecasts with MSPE lower than that for the random walk model for any currency. Consequently, using unadjusted statistics with standard normal critical values provides no evidence of exchange rate predictability for any currency in our sample. The p-values for DMW statistic calculated using bootstrap do not provide significant evidence of predictability either.

UIRP Fundamentals

We estimate the model with UIRP fundamentals for all countries except Denmark, Netherlands and Portugal, which are excluded from the sample because of the lack of interest rate data. This leaves 9 countries in the sample. Using the CW statistic, the model with UIRP fundamentals significantly outperforms the random walk for 3 out of 9 countries (Canada at 1% level, and Japan and U.K at 10 % level). Clark and West (2005a) examine the predictability of UIRP fundamentals for 4 countries: Japan, Switzerland, Canada and the U.K. Looking at a slightly shorter period of recent data they obtain similar results. [21] They find that the UIRP fundamental-based model significantly outpredicts the random walk for Canada and Switzerland (at the 5% level). Using more data we find that the UIRP fundamentals perform better for the U.K. However, since we

[21] Clark and West (2005a) examine the period from January 1980 to October 2003 for Canada and Japan, and from January 1975 to October 2003 for Switzerland and the U.K.

are using a shorter period of data for Switzerland, which starts in January 1980 instead of January 1975, we are unable to detect statistically significant superior performance of the model for the Swiss franc. Overall, the evidence of exchange rate predictability with UIRP fundamentals is about the same as with monetary fundamentals. Once again, we find no evidence of exchange rate predictability using the DMW statistic with standard normal critical values. According to the p-values calculated using bootstrap for the DMW statistic, the model with UIRP fundamentals significantly outperforms the random walk only for Canada (at the 1% level).

PPP Fundamentals

For the PPP model, using CW procedure makes an adjusted MSPE of the model smaller than that of the random walk for 8 out of 12 countries. However, the model significantly outperforms the random walk only for Portugal (at 5% level) and Denmark (at 10% level). As with the monetary model, no evidence of exchange rate predictability is found using the DMW statistic with either normal or bootstrapped critical values.

The evidence of exchange rate predictability with monetary, PPP, and UIRP fundamentals is clearly weaker than that with Taylor rule fundamentals. Using the CW statistic, none of the models significantly outperforms the random walk for more than 3 countries and, and combining the three models, evidence of predictability can be found for 6 of the 12 countries, compared with 8 of the 12 countries using the Taylor rule models. Using the DMW statistic with bootstrapped critical values and combining the three models, evidence of predictability can be found for only 1 of the 12 countries, compared with 6 of the 12 countries using the Taylor rule models.

4.2 Long Horizon Forecasts

For longer horizons (from 3 to 36 month), we construct the test statistic as described in section 3.1 of Clark and West (2005a). Extending the forecast horizon beyond h=1 introduces serial correlation of order (h-1) in the error term. With overlapping data the resulting time series $2y_{t+h,h}X'_{t+1}\hat{\beta}_t$ follows a MA(h-1) process under the null. Thus, we need to account for serial correlation when performing inference for horizons h>1. Clark and West (2005a) suggest using the heteroskedasticity and autocorrelation consistent (HAC) estimator developed in West (1996), which in our case reduces to a mean adjusted version of the Hodrick (1992) estimator. Table 4 summarizes the results for long-horizon forecasts. The Taylor rule fundamentals are constructed using the deviation of actual output from its quadratic trend as a measure of the output gap. According to Clark and West (2005a), using the West-Hodrick HAC estimator in long-horizon tests results in more accurately sized tests than using the Andrews (1991) or Andrews-Monahan (1991) HAC estimator.[22] The p-values for the DMW statistic are calculated using bootstrap.

[22]The results for the test statistics with Andrews (1991) HAC estimator exhibits a similar pattern.

Table 4 shows that the out-of-sample predictability of the exchange rate models does not increase significantly as the horizon increases for any model or currency. Using the new inference procedure we do not find evidence supporting increased forecasting ability in the long run. Although the p-values decrease at very long horizons of 36 months for the majority of models and currencies, the decrease is never enough to indicate exchange rate predictability of the model. None of the models predicts the exchange rate significantly better than the random walk at horizons longer than 6 months.

Engel and West (2005) argue that when fundamentals are I(1) and the factor for discounting future fundamentals is close to one, the present value asset pricing model, which is consistent with a wide range of exchange rate models including those considered in this paper, places greater weight on future fundamentals than on current fundamentals, which therefore have weak forecasting power. Wang (2006) shows that, under the Engel-West explanation, long-horizon regressions have more predictive ability than the short-horizon regressions only when the change of fundamentals is highly persistent, and, even in this case, this increase is generally small and hard to detect in finite samples. Our results confirm the difficulty of detecting higher predictability in long-horizon regressions and provide some empirical support for the findings in Wang (2006). He shows, both analytically and using Monte Carlo methods, that under the moderate amount of persistence of the change of fundamentals actually observed in the data, the change of the exchange rate is less serially correlated in long-horizon data than in short-horizon data. Therefore, the exchange rate behaves more like a random walk in long-horizon than in short-horizon data, making it harder to find statistical evidence in favor of the economic model over the random walk in long-horizon regressions than in short-horizons regressions.

5. Conclusions

Research on exchange rate predictability has come full circle from the "no predictability at short horizons" results of Meese and Rogoff (1983a, 1983b) to the "predictability at long horizons but not short horizons" results of Mark (1995) and Chen and Mark (1996) to the "no predictability at any horizons" results of Cheung, Chinn, and Pascual (2005). We come to a very different conclusion, reporting strong evidence of exchange rate predictability at the one-month horizon, slight evidence of predictability at the six-month horizon, and no evidence of predictability at longer horizons. [23]

We show that model selection and inference methodology both play an important role in providing evidence of exchange rate predictability. Using the CW statistic, which adjusts for bias in the MSPE from the alternative model under the null, we provide evidence of short-run exchange rate predictability for models with Taylor rule fundamentals and 1970s vintage models with monetary, UIRP, and PPP fundamentals. Had we used the unadjusted DMW statistic with standard normal critical values, we would have incorrectly

[23] Using a new bootstrap technique and/or extending the sample size used in Mark (1995), Kilian (1999) finds evidence that the forecasting ability of the monetary exchange rate model decreases or stays the same with the horizon.

concluded that none of the models could forecast better than a random walk at any horizon even when the economic model actually had statistically significant predictive power.

While we find some evidence of short-run predictability using models with monetary, UIRP, and PPP fundamentals, the Taylor rule models perform better out-of sample with the model that incorporates exchange rate targeting in the foreign country having the best forecasting ability. At the one-month-ahead horizon, we find evidence of predictability using the Taylor rule models for 8 out of 12 currencies. We do not find statistical evidence that exchange rate predictability increases with the horizon, and we find no evidence of exchange rate predictability with economic models at any horizons beyond six months.

Our results can be interpreted in the context of Engel and West (2005), who show that, if fundamentals are I(1) and the discount factor is large, exchange rates will approximately follow a random walk and it is not surprising that present-value models do not produce evidence of predictability. Wang (2006) shows that, under the moderate degree of persistence of the change of fundamentals observed in the data, short-horizon regressions have more power than long-horizon regressions. We demonstrate the importance of inference methodology in assessing these results. Using the (undersized) DMW statistic with standard normal critical values, we "confirm" previous work that fails to find evidence of predictability at any horizon with either conventional or Taylor rule models. It is only by using the (correctly sized) CW statistic, especially in conjunction with the Taylor rule models, that we find evidence of predictability at short-run, but not long-run, horizons.

References

Alquist, Ron and Menzie Chinn, "Conventional and Unconventional Approaches to Exchange Rate Modeling and Assessment," NBER Working Paper No.12481, 2006

Andrews, Donald W. K.., "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica*, 1991, 59, pp. 817-858.

Andrews, Donald W. K.. and J. Christopher Monahan, "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator." *Econometrica*, 1991, 60, pp. 953-966.

Berkowitz, Jeremy and Lorenzo Giorgianni, "Long-Horizon Exchange Rate Predictability?" *Review of Economics and Statistics*, February 2001, 83(1), pp. 81-91

Benjamini, Yoav and Yosef Hochberg, "Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing" *Journal of Royal Statistical Society*, 1995, Series B, 57(1), pp.289-300

Chen, Jian and Nelson Mark, "Alternative Long-Horizon Exchange Rate Predictors" *International Journal of Finance and Economics,* October 1996, 1, pp.229-250

Cheung, Yin-Wong; Menzie D. Chinn, and Antonio Garcia Pascual, "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?" *Journal of International Money and Finance*, 2005, 24, 1150-1175.

Chinn, Menzie D., "The (Partial) Rehabilitation of Interest Rate Parity in the Floating Rate Era: Longer Horizons, Alternative Expectations and Emerging Markets" *Journal of International Money and Finance*, 2006, 26, pp.7-21

Chinn, Menzie D. and Richard A. Meese, "Banking on Currency Forecasts: How Predictable Is Change in Money*?" Journal of International Economics*, February 1995, 38(1-2), pp. 161-178

Chinn, Menzie D. and Guy Meredith, "Monetary Policy and Long-Horizon Uncovered Interest Parity." *IMF Staff Papers* 51(3), 2004, pp. 409-430.

Clarida, Richard, Jordi Gali, and Mark Gertler,. "Monetary Rules in Practice: Some International Evidence." *European Economic Review*, 1998, 42, pp. 1033-1067

Clark, Todd E. and Michael W. McCracken, "Tests of Equal Forecast Accuracy and Encompassing for Nested Models" *Journal of Econometrics*, 2001, 105, pp. 671-110

_______ , "The Predictive Content of the Output Gap for Inflation" manuscript, *Journal of Money, Credit and Banking*, forthcoming, 2005a

_______ , "Evaluating Direct Multi-Step Forecasts", *Econometric Reviews*, 2005b, 24, pp. 369-404

Clark, Todd and Kenneth West, "Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis." *Journal of Econometrics*, forthcoming, 2005a

_______ , "Approximately Normal Tests For Equal Predictive Accuracy in Nested Models" *Journal of Econometrics,* forthcoming, 2005b

Corradi, Valentina and Norman R. Swanson, "Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes" *International Economic Review*, forthcoming, 2005

Diebold, Francis and Robert Mariano, "Comparing Predictive Accuracy" *Journal of Business and Economic Statistics*, 1995, 13, 253-263

Engel, Charles and Kenneth West, "Exchange Rate and Fundamentals," *Journal of Political Economy*, 2005, 113, 485-517

_______ , "Taylor Rules and the Deutschmark-Dollar Real Exchange Rates" *Journal of Money, Credit and Banking*, 2006, 38, 1175-1194.

Faust, Jon R., John Rogers and Jonathan H. Wright, "Exchange Rate Forecasting: The Errors We've Really Made." *Journal of International Economics*, 2003, 60, 35-59

Gourinchas, Pierre-Olivier and Helene Rey, "International Financial Adjustment," manuscript, Princeton University, 2005

Hodrick, Robert J., "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *The Review of Financial Studies*, 1992, 5(3), pp. 357-386

Hodrick, Robert J. and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation" *Journal of Money, Credit, and Banking*, 1997, 29, pp. 1-16

Inoue, Atsushi, and Lutz Kilian, "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?" *Econometric Reviews*, 2004, 23, 371-402.

Kilian, Lutz, "Exchange Rates and Monetary Fundamentals: What Do We Learn From Long-Horizon Regressions?" *Journal of Applied Econometrics*, 1999, 14(5), pp. 491-510

Mark, Nelson, "Exchange Rate and Fundamentals: Evidence on Long-Horizon Predictability" *American Economic Review*, March 1995, 85(1), pp.201-218

_______, "Changing Monetary Policy Rules, Learning and Real Exchange Rate Dynamics" NBER Working Paper 11061, 2005

Mark, Nelson and Donggyu Sul, "Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel." *Journal of International Economics*, 2001, 53, pp. 29-52

McCracken, Michael W. "Asymptotics for Out-Of-Sample Tests of Causality", *Journal of Economerics*, forthcoming, 2006

McCracken, Michael W. and Stephen Sapp, "Evaluating the Predictability of Exchange Rates Using Long Horizon Regressions: Mind Your p's and q's!", *Journal of Money, Credit and Banking,* 2005, 37, pp. 473-494

Meese, Richard A. and Kenneth Rogoff, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample*?" Journal of International Economics*, February 1983a, 14(1-2), pp. 3-24

_______. "The Out of Sample Failure of Empirical Exchange Rate Models" in Jacob A. Frenkel, ed., *Exchange Rates and International Macroeconomics*. Chicago, IL: University of Chicago Press, 1983b, pp.6-105

Rossi, Barbara, "Testing Long-Horizon Predictive Ability with High Persistence, and the Meese-Rogoff Puzzle", *International Economic Review*, February 2005, 46(1), pp. 61-92

Storey, John D., "A Direct Approach to False Discovery Rates" *Journal of Royal Statistical Society*, 2002, Series B, 64(3), pp.479-498

_______., "The Positive False Discovery Rate: A Bayesian Interpretation and the q-value" *The Annals of Statistics*, 2003, 31(6), pp.2013-2035

Storey, John D. and Robert Tibshirani, "Statistical Significance for Genomewide Studies" *Proceedings of the National Academy of Sciences*, 2003, 100, pp.9440-9445

Papell, David H., "The Panel Purchasing Power Parity Puzzle" *Journal of Money, Credit, and Banking*, March 2006, 38(2), pp. 447-467

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

Wang, Jiang, "Can Long Horizon Data Beat Random Walk Under Engel-West Explanation?", mimeo University of Wisconsin-Maddison, 2006

West, Kenneth D., "Asymptotic Inference about Predictive Ability" *Econometrica*, 1996, 64, pp.1067-1084

_______. "Forecast Evaluation" manuscript, mimeo University of Wisconsin, 2005

Table 1. p and q-values for Symmetric Taylor Rule Model

Country		*Linear Trend*			*Quadratic Trend*			*HP Filter*		
		DMW standard	*DMW bootstrap*	*CW*	*DMW standard*	*DMW bootstrap*	*CW*	*DMW standard*	*DMW bootstrap*	*CW*
Japan										
	p-value	0.745	0.242	0.074*	0.938	0.607	0.166	0.662	0.272	0.089*
	q-value		0.362	0.252		0.489	0.344		0.377	0.272
Canada										
	p-value	0.177	0.004***	0.006***	0.258	0.007***	0.008***	0.154	0.006***	0.006***
	q-value		0.098*	0.098*		0.098*	0.098*		0.098*	0.098*
Switzerland										
	p-value	0.922	0.681	0.575	0.899	0.578	0.524	0.745	0.390	0.295
	q-value		0.500	0.486		0.486	0.479		0.430	0.382
U.K.										
	p-value	0.773	0.166	0.271	0.735	0.143	0.218	0.417	0.030**	0.035**
	q-value		0.344	0.377		0.323	0.348		0.181	0.188
France										
	p-value	0.773	0.183	0.195	0.773	0.214	0.206	0.564	0.081*	0.080*
	q-value		0.345	0.345		0.347	0.348		0.261	0.261
Germany										
	p-value	0.892	0.490	0.567	0.843	0.403	0.433	0.722	0.262	0.224
	q-value		0.466	0.486		0.429	0.441		0.377	0.348
Italy										
	p-value	0.920	0.825	0.136	0.802	0.667	0.042**	0.575	0.377	0.025**
	q-value		0.533	0.323		0.500	0.188		0.430	0.181
Sweden										
	p-value	0.964	0.675	0.348	0.944	0.565	0.288	0.938	0.619	0.599
	q-value		0.500	0.406		0.486	0.380		0.493	0.489
Australia										
	p-value	0.855	0.276	0.252	0.824	0.382	0.484	0.984	0.928	0.575
	q-value		0.377	0.368		0.429	0.465		0.561	0.486
Denmark										
	p-value	0.752	0.184	0.192	0.846	0.323	0.319	0.995	0.941	0.930
	q-value		0.345	0.345		0.388	0.387		0.563	0.561
Netherlands										
	p-value	0.957	0.783	0.674	0.946	0.749	0.691	0.955	0.806	0.810
	q-value		0.533	0.500		0.524	0.500		0.533	0.533
Portugal										
	p-value	0.755	0.897	0.128	0.841	0.935	0.145	0.548	0.744	0.038**
	q-value		0.555	0.323		0.562	0.323		0.524	0.188

Notes to Tables 1 and 2:

1. The models are estimated using data from March, 1973 for all countries except Denmark, for which output is only available from January 1974.
2. To detrend the monthly output series using HP filter, we use the value of smoothing parameter equal to 14400.
3. The output gap in period t is calculated using all the currently available data. The output gap for the first period is calculated using output series from 1971:1 to 1973:3.
4. The p-values in columns DMW standard and CW are based on standard normal critical values. The p-values in columns DMW bootstrap are calculated using bootstrap. The q-values are calculated using the algorithm due to Storey (2003).
5. In all of the tables, *, **, and *** denote test statistics significant at 10, 5, and 1% level, respectively, based on critical values for one-sided tests.

Table 2. p- and q-values for Taylor Rule Model with Exchange Rate Targeting

Country		*Linear Trend*			*Quadratic Trend*			*HP Filter*		
		DMW standard	*DMW bootstrap*	*CW*	*DMW standard*	*DMW bootstrap*	*CW*	*DMW standard*	*DMW bootstrap*	*CW*
Japan										
	p-value	0.972	0.566	0.140	0.954	0.477	0.130	0.943	0.494	0.334
	q-value		0.486	0.323		0.464	0.323		0.466	0.397
Canada										
	p-value	0.663	0.043**	0.053*	0.504	0.017**	0.045**	0.401	0.013**	0.040**
	q-value		0.188	0.216		0.170	0.190		0.145	0.188
Switzerland										
	p-value	0.992	0.865	0.816	0.968	0.678	0.591	0.864	0.393	0.285
	q-value		0.546	0.533		0.500	0.486		0.430	0.380
U.K.										
	p-value	0.785	0.119	0.136	0.788	0.111	0.125	0.429	0.018**	0.005***
	q-value		0.323	0.323		0.309	0.323		0.170	0.098*
France										
	p-value	0.890	0.266	0.143	0.901	0.296	0.171	0.853	0.233	0.097*
	q-value		0.377	0.323		0.382	0.344		0.352	0.277
Germany										
	p-value	0.988	0.843	0.788	0.973	0.720	0.610	0.957	0.657	0.560
	q-value		0.535	0.533		0.513	0.489		0.500	0.486
Italy										
	p-value	0.964	0.819	0.193	0.883	0.589	0.065*	0.839	0.521	0.043**
	q-value		0.533	0.345		0.486	0.241		0.479	0.188
Sweden										
	p-value	0.980	0.573	0.067*	0.956	0.436	0.061*	0.975	0.584	0.164
	q-value		0.486	0.242		0.441	0.234		0.486	0.344
Australia										
	p-value	0.885	0.224	0.681	0.945	0.408	0.758	0.985	0.886	0.732
	q-value		0.347	0.500		0.431	0.528		0.551	0.519
Denmark										
	p-value	0.776	0.091*	0.031**	0.812	0.127	0.037**	0.977	0.642	0.309
	q-value		0.272	0.181		0.323	0.188		0.500	0.383
Netherlands										
	p-value	0.958	0.575	0.425	0.964	0.631	0.508	0.977	0.764	0.652
	q-value		0.486	0.441		0.499	0.472		0.529	0.500
Portugal										
	p-value	0.648	0.640	0.024**	0.716	0.694	0.027**	0.437	0.439	0.008***
	q-value		0.500	0.181		0.500	0.181		0.441	0.098*

Table 3. p- and q-values for the UIRP, PPP, and Monetary Models

Country		*UIRP Model*			*PPP Model*			*Monetary Model (k=0)*		
		DMW standard	*DMW bootstrap*	*CW*	*DMW standard*	*DMW bootstrap*	*CW*	*DMW standard*	*DMW bootstrap*	*CW*
Japan										
	p-value	0.688	0.443	0.061*	0.941	0.672	0.605	0.680	0.175	0.028**
	q-value		0.441	0.234		0.500	0.488		0.345	0.181
Canada										
	p-value	0.149	0.005***	0.003***	0.965	0.779	0.824	0.829	0.277	0.209
	q-value		0.098*	0.098*		0.533	0.533		0.377	0.348
Switzerland										
	p-value	0.844	0.457	0.179	0.941	0.779	0.691	0.985	0.400	0.022**
	q-value		0.452	0.345		0.533	0.500		0.429	0.181
U.K.										
	p-value	0.702	0.217	0.070*	0.695	0.230	0.382	0.722	0.200	0.316
	q-value		0.348	0.245		0.352	0.430		0.345	0.387
France										
	p-value	0.977	0.876	0.831	0.874	0.435	0.341	0.998	0.986	0.961
	q-value		0.550	0.533		0.441	0.402		0.580	0.569
Germany										
	p-value	0.831	0.466	0.198	0.866	0.570	0.488	0.979	0.880	0.839
	q-value		0.457	0.345		0.486	0.466		0.550	0.535
Italy										
	p-value	0.749	0.537	0.309	0.959	0.919	0.500	0.995	0.990	0.903
	q-value		0.486	0.383		0.560	0.468		0.581	0.556
Sweden										
	p-value	0.999	0.998	0.949	0.966	0.709	0.152	0.844	0.422	0.252
	q-value		0.582	0.564		0.508	0.333		0.441	0.368
Australia										
	p-value	0.758	0.585	0.192	0.971	0.825	0.829	0.855	0.374	0.305
	q-value		0.486	0.345		0.533	0.533		0.429	0.382
Denmark										
	p-value	-	-	-	0.810	0.222	0.087*	0.821	0.170	0.094*
	q-value					0.347	0.272		0.344	0.274
Netherlands										
	p-value	-	-	-	0.770	0.391	0.309	0.776	0.283	0.398
	q-value					0.429	0.383		0.380	0.430
Portugal										
	p-value	-	-	-	0.591	0.830	0.029**	0.903	0.907	0.648
	q-value					0.533	0.181		0.556	0.500

Notes:

1. The UIRP model is estimated using data from January 1980 for Japan and Switzerland, March 1977 for Italy, and July 1975 for Germany because the interest rate data are not available prior to those periods. Also, the Treasury bill series stops in October 2003 for Japan, March 2004 for Sweden, and June 2004 for Australia. The monetary model is estimated using the data from December 1977 for France, December 1974 for Italy, and December 1979 for Portugal because of the lack of money supply data prior to those periods.
2. The p-values in columns DMW standard and CW are based on standard normal critical values. The p-values in columns DMW bootstrap are calculated using bootstrap. The q-values are calculated using the algorithm due to Storey (2003).

Table 4. p-values for DMW and CW Statistics: h-period Ahead Forecasts

Horizon (h)	*UIRP Model*		*PPP Model*		*Monetary Model (k=0)*		*Taylor Rule Model (symmetric)*		*Taylor Rule Model (asymmetric)*	
	A.Japan									
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.443	0.061	0.672	0.605	0.175	0.028	0.607	0.166	0.477	0.130
6	0.427	0.255	0.933	0.625	0.540	0.356	0.852	0.440	0.745	0.484
12	0.747	0.421	0.948	0.595	0.752	0.484	0.922	0.484	0.678	0.488
36	0.143	0.409	0.965	0.548	0.831	0.512	0.767	0.512	0.695	0.532
	B.Canada									
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.005	0.003	0.779	0.824	0.277	0.209	0.007	0.008	0.017	0.045
6	0.079	0.245	0.884	0.674	0.712	0.504	0.052	0.230	0.130	0.323
12	0.408	0.498	0.996	0.698	0.905	0.625	0.337	0.440	0.627	0.576
36	0.648	0.501	0.972	0.556	0.999	0.618	0.517	0.502	0.512	0.516
	C.Switzerland									
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.457	0.179	0.779	0.691	0.400	0.022	0.578	0.524	0.678	0.591
6	0.782	0.484	0.801	0.564	0.683	0.312	0.927	0.633	0.901	0.625
12	0.968	0.567	0.941	0.587	0.844	0.440	0.992	0.663	0.991	0.674
36	0.996	0.536	0.702	0.508	0.792	0.466	0.367	0.476	0.841	0.499
	D.United Kingdom									
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.217	0.070	0.230	0.382	0.200	0.316	0.143	0.218	0.111	0.125
6	0.381	0.345	0.520	0.532	0.617	0.544	0.557	0.468	0.606	0.444
12	0.525	0.440	0.611	0.516	0.851	0.544	0.672	0.516	0.560	0.464
36	0.924	0.516	0.727	0.504	0.971	0.556	0.951	0.524	0.993	0.548
	E.France									
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.876	0.831	0.435	0.341	0.986	0.961	0.214	0.206	0.296	0.171
6	0.965	0.715	0.859	0.618	0.963	0.691	0.862	0.629	0.800	0.571
12	0.993	0.662	0.992	0.663	0.978	0.640	0.882	0.587	0.922	0.603
36	0.871	0.663	0.977	0.567	0.999	0.640	0.120	0.464	0.598	0.498
	F.Germany									
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.466	0.198	0.570	0.488	0.880	0.839	0.403	0.433	0.720	0.610
6	0.808	0.468	0.686	0.532	0.809	0.579	0.644	0.552	0.869	0.575
12	0.920	0.498	0.884	0.571	0.794	0.528	0.871	0.576	0.979	0.603
36	0.999	0.498	0.941	0.520	0.639	0.496	0.612	0.492	0.985	0.524

Table 4 (Continued)

Horizon (h)	*UIRP Model*		*PPP Model*		*Monetary Model (k=0)*		*Taylor Rule Model (symmetric)*		*Taylor Rule Model (asymmetric)*	
					G.Italy					
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.537	0.309	0.919	0.500	0.990	0.903	0.667	0.042	0.589	0.065
6	0.716	0.492	0.994	0.677	0.998	0.794	0.914	0.413	0.826	0.472
12	0.647	0.464	0.997	0.629	0.996	0.655	0.950	0.492	0.870	0.524
36	0.945	0.520	0.998	0.579	0.999	0.534	0.979	0.524	0.832	0.528
					H.Sweden					
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.998	0.949	0.709	0.152	0.422	0.252	0.565	0.288	0.436	0.061
6	0.998	0.857	0.954	0.528	0.824	0.540	0.974	0.695	0.950	0.552
12	0.999	0.757	0.989	0.583	0.973	0.583	0.990	0.652	0.981	0.587
36	0.999	0.603	0.861	0.536	0.946	0.528	0.992	0.552	0.629	0.520
					I.Australia					
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	0.585	0.192	0.825	0.829	0.374	0.305	0.382	0.484	0.408	0.758
6	0.750	0.413	0.955	0.719	0.505	0.444	0.991	0.409	0.995	0.575
12	0.742	0.492	0.969	0.640	0.857	0.556	0.971	0.484	0.971	0.603
36	0.987	0.958	0.929	0.560	0.838	0.536	0.492	0.548	0.211	0.524
					K.Denmark					
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	-	-	0.222	0.087	0.170	0.094	0.323	0.319	0.127	0.037
6	-	-	0.631	0.448	0.812	0.512	0.663	0.524	0.344	0.337
12	-	-	0.954	0.531	0.975	0.583	0.716	0.520	0.652	0.460
36	-	-	0.997	0.603	0.963	0.540	0.273	0.476	0.395	0.488
					L.Netherlands					
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	-	-	0.391	0.309	0.283	0.398	0.749	0.691	0.631	0.508
6	-	-	0.727	0.540	0.706	0.595	0.982	0.734	0.871	0.652
12	-	-	0.899	0.571	0.917	0.606	0.997	0.691	0.965	0.633
36	-	-	0.971	0.536	0.734	0.520	0.967	0.564	0.999	0.625
					M.Portugal					
	DMW	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*	*DMW*	*CW*
1	-	-	0.830	0.029	0.907	0.648	0.935	0.145	0.694	0.027
6	-	-	0.985	0.417	0.924	0.603	0.995	0.560	0.873	0.456
12	-	-	0.996	0.508	0.962	0.599	0.999	0.610	0.932	0.536
36	-	-	0.999	0.548	0.975	0.579	0.993	0.528	0.961	0.560

Notes :

1. The Taylor rule models use quadratic trend of industrial production series and a measure of potential output.
2. The p-values for DMW statistic are calculated using bootstrap.