BANK OF CANADA
BANQUE DU CANADA

Discussion Paper/Document d'analyse
2011-8

Exchange Rates and Individual Good's Price Misalignment: Some Preliminary Evidence of Long-Horizon Predictability

by Wei Dong and Deokwoo Nam

www.bank-banque-canada.ca

Bank of Canada Discussion Paper 2011-8

October 2011

Exchange Rates and Individual Good's Price Misalignment: Some Preliminary Evidence of Long-Horizon Predictability

by

Wei Dong[1] and Deokwoo Nam[2]

[1]Canadian Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
wdong@bankofcanada.ca

[2]City University of Hong Kong
Hong Kong
deokwnam@city.edu.hk

ISSN 1914-0568

Acknowledgements

We are indebted to Charles Engel for motivating the idea of this work and for his comments and suggestions at various stages of this paper. We also thank Greg Bauer, Mick Devereux, Nelson Mark, Lise Pichette and seminar participants for helpful comments.

Abstract

When prices are sticky, movements in the nominal exchange rate have a direct impact on international relative prices. A relative price misalignment would trigger an adjustment in consumption and employment, and may help to predict future movements in the exchange rate. Although purchasing-power-parity fundamentals, in general, have only weak predictability, currency misalignment may be indicated by price differentials for some goods, which could then have predictive power for subsequent re-evaluation of the nominal exchange rate. The authors collect good-level price data to construct deviations from the law of one price and examine the resulting price-misalignment model's predictive power for the nominal exchange rates between the U.S. dollar and two other currencies: the Japanese yen and the U.K. pound. To account for small-sample bias and data-mining issues, inference is drawn from bootstrap distributions and tests of superior predictive ability (SPA) are performed. The slope coefficients and R-squares increase with the forecast horizon for the bilateral exchange rates between the U.S. dollar and the Japanese yen and the U.S. dollar and the U.K. pound. The out-of-sample SPA tests suggest that the authors' price-misalignment model outperforms random walks either with or without drift for the U.S. dollar vis-à-vis the Japanese yen at the 5 per cent level of significance over long horizons.

JEL classification: F31, F47
Bank classification: Exchange rates; International topics

Résumé

En situation de rigidité des prix, les variations du taux de change nominal ont des incidences directes sur les prix relatifs entre pays. Un déséquilibre des prix relatifs déclenche un ajustement de la consommation et de l'emploi, et apporte un éclairage susceptible de faciliter la prévision de l'évolution future du taux de change. Bien que le modèle reposant sur la théorie de la parité des pouvoirs d'achat n'ait souvent qu'un faible pouvoir prédictif, il arrive que les différences de prix entre certains biens soient l'indice de déséquilibres du taux de change; ces différences pourraient alors servir à prévoir l'ampleur de la réévaluation à venir du taux de change nominal. Les auteurs recueillent des données sur les prix des biens de manière à en dégager les écarts par rapport à la loi du prix unique, puis examinent si leur modèle de déséquilibre des prix permet de prévoir le taux de change nominal entre le dollar américain et deux autres monnaies : le yen et la livre sterling. Les auteurs parent au biais créé par la petite taille de l'échantillon et aux problèmes de surexploitation des données en tirant leurs déductions de distributions découlant d'une procédure *bootstrap* et en faisant appel au test de « supériorité prédictive » ou test SPA (pour *superior predictive ability*) que propose Hansen. Le coefficient et le R^2 augmentent avec l'horizon de projection des deux taux de change bilatéraux étudiés ($ É.-U./¥ et $ É.-U./£). Les tests SPA réalisés hors échantillon donnent à penser qu'aux horizons éloignés, le modèle proposé de déséquilibre des prix surclasse les modèles de marche aléatoire (avec dérive ou non) au seuil de 5 % dans le cas du taux de change $ É.-U./¥.

Classification JEL : F31, F47
Classification de la Banque : Taux de change; Questions internationales

1 Introduction

Understanding the connection between exchange rates and macroeconomic fundamentals has been one of the central challenges in international macroeconomics since the early 1970s. Although exchange rates are highly volatile, they should reflect basic macroeconomic fundamentals such as interest rates, purchasing power and trade balances. As such, international economists have long held out hope that they could explain exchange rates with these fundamentals. Unfortunately, in practice, the performance of structural exchange rate models has not been very satisfying. As first shown by Meese and Rogoff (1983), such models can hardly beat a random walk process when it comes to out-of-sample forecasting.

Recently, some studies have found certain forecasting power in monetary models at horizons of two to four years (Mark 1995; Engel, Mark and West 2007). Other attempts to forecast at more policy-relevant shorter horizons have also reported positive forecasting results (Gourinchas and Rey 2007; Engel, Mark and West 2007; Molodtsova and Papell 2009). However, when these positive results are re-examined for their econometric approaches and alternative time windows, they do not hold over both long horizons (Kilian 1999; Berkowitz and Giorgianni 2001) and short horizons (Rogoff and Stavrakeva 2008). As a result, despite notable methodological improvements, we are not much closer to being able to forecast exchange rates. In this paper, we examine out-of-sample exchange rate predictability based on good-level price deviations from the law of one price (LOOP) and conclude that, after accounting for the econometric concerns of the bootstrap method and small-sample bias, there is some evidence of exchange rate predictability with disaggregate price-misalignment fundamentals over long horizons.

Most exchange rate movements in the short run seem to reflect changes in expectations about future monetary or real conditions. When prices are sticky, however, movements in the nominal exchange rate have a direct impact on international relative prices. For some goods, the relative prices across countries should reflect the current levels of their demand and supply, rather than future expectations, and therefore changes in the nominal exchange rate may have undesirable allocation effects. In other words, when the changes in the exchange rate are primarily forward looking, the relative prices would be forced to incorporate these expectation effects, and the terms of trade or other international prices may be badly misaligned in the short run. The relative price misalignment caused by the dual role of the exchange rate in goods and asset markets would trigger adjustment in consumption and employment. For example, if the prices of certain goods are more expensive in Japan than in the United States, consumers in both Japan and the United States would prefer to purchase more U.S. goods. The increased demand for U.S. goods will drive the U.S. dollar to appreciate with respect to the Japanese yen. In the absence of transportation and other transactions costs, competitive markets will equalize the price of an identical good in two countries when the prices are expressed in the same currency. Therefore, purchasing-power parity (PPP) may serve as an anchor for long-run real exchange rates, and price misalignments may help to predict subsequent re-evaluation of the nominal exchange rate. Based on this reasoning, models of PPP naturally lend themselves to determining whether a currency is overvalued or undervalued. However, when it comes to forecasting future movements in the exchange rate, many models outperform PPP in terms of exchange rate predictability; for example, monetary models and Taylor rule models. This may be because, although real exchange rates may converge to parity in the long run, the rate at which this happens is so slow that it is at best of little practical relevance over horizons of concern to the forecasting of exchange rates.

Although it may seem as if PPP and the LOOP are the same, they do in fact differ: PPP applies to the aggregate price level and the LOOP applies to individual goods. First, there is a great amount of heterogeneity for prices at the good level. In fact, it is hard to think of reasons why all relative prices comprising the consumption-basket real exchange rate should converge to parity at the same speed. Failure to allow for the heterogeneity in price-adjustment dynamics at the good level may even induce a positive aggregation bias in persistence estimates (Imbs et al. 2005). Thus, although PPP as a fundamental does not generally work well in predicting exchange rates, the LOOP for some goods may. Second, for certain goods, currency misalignments may be indicated by price deviations, which could further help to predict future movements in the exchange rate. For some other goods, their spot prices may also reflect both current and future supply and demand conditions; for example, oil. In that case, both the price differentials and exchange rates are subject to the same set of shocks, and it is hard to point to the direction of forecasting. This may partially account for the weak predictive power of price deviations at the aggregate level. This paper therefore examines whether *disaggregated* price misalignments have any predictive power for future movements in the exchange rate.

Specifically, we examine monthly observations of nominal exchange rates between the U.S. dollar and the Japanese yen from 1973M03 to 2009M08, and the U.S. dollar and the U.K. pound from 1987M01 to 2009M08. We collect good-level price data for 67 items in the U.S.-Japan case and 48 items in the U.S.-U.K. case to construct deviations from the LOOP. We define price misalignment as the deviation from the LOOP for a certain good. With the price-misalignment series, we then perform in-sample and out-of-sample forecasting analysis for changes in the nominal exchange rate. We emphasize three aspects of our econometric approach.

First, we use the Clark and West tests based on Newey-West heteroskedasticity and autocorrelation consistent (HAC) estimates. The relevant literature on exchange rate predictability compares the out-of-sample predictability of models on the basis of different measures. The most commonly used measure of predictive ability is the mean squared prediction error (MSPE). Diebold and Mariano (1995) tests are often used to evaluate the out-of-sample performance of the models based on the MSPE comparison. While the Diebold and Mariano (DM) tests are appropriate for non-nested models, they are asymptotically invalid when testing nested models (Clark and McCracken 2001; Corradi and Swanson 2004, 2007). Thus, alternatively, we evaluate the equal forecast accuracy using Clark and West's (2007) procedure. We adopt Newey-West's (1987) HAC covariance matrix estimator with Andrews's (1991) procedure for selecting a truncation lag, so as to account for serial correlation when the forecast horizon is more than one period.

Second, in our bootstrap algorithm, we take into account the methodological suggestions made by Kilian (1999) to achieve consistency in the test procedure and correct for small-sample bias. Moreover, as a robustness check, we also conduct bootstrap analysis under the restricted vector error correction model (VECM).

Finally, to account for data-mining issues, inference is drawn from tests of superior predictive ability (SPA). We are comparing the benchmark random walks to a possibly large set of candidate models; in such a situation, a few pairwise tests can signal dominance of one model over the other simply by chance and lead to the rejection of the null hypothesis. To address this data-snooping problem, we apply the SPA test proposed by Hansen (2005) and based on the seminal paper by White (2000).

We summarize our main findings here. First, for the bilateral exchange rates between the U.S. dollar and the Japanese yen and the U.S. dollar and the U.K. pound, we find that the slope coefficients and R-squares from price-misalignment models increase with the forecast horizon. Second, for the U.S. dollar-Japanese yen exchange rate, the out-of-sample SPA test results suggest that our price-misalignment model outperforms random walks either with or without drift at the 5 per cent level of significance over long horizons (12 months). Third, price deviations on electricity and frozen fish and seafood can predict the bilateral exchange rate between the U.S. dollar and the Japanese yen both in-sample and out-of-sample at almost all forecast horizons.

The remainder of this paper is organized as follows. Section 2 motivates the price-misalignment model that we estimate. The data are described in section 3. Section 4 describes the empirical methodology including the bootstrap procedure used to conduct inferences and the SPA test. The empirical results are reported in section 5. Section 6 offers some conclusions.

2 Motivating the Price-Misalignment Model

The law of one price in its absolute version may be written as:

$$P_{i,t} = S_t P^*_{i,t}, \tag{1}$$

where $P_{i,t}$ denotes the price of good i in the home country (U.S.) denominated in terms of domestic currency, $P^*_{i,t}$ denotes the price of good i in the foreign country denominated in foreign currency, and S_t represents the nominal exchange rate defined as the U.S. dollar per foreign currency. In theory, the LOOP should hold, based on the idea of frictionless good arbitrage. However, there are three caveats empirically: (i) transportation costs, barriers to trade, and other transactions costs can be significant; (ii) there must be competitive markets for the goods and services in both countries for the LOOP to hold; and (iii) the LOOP applies only to tradable goods — immobile goods, such as many services that are local, are of course not traded between countries. In fact, econometric studies suggest rejection of the LOOP for a very broad range of goods and provide empirical evidence that deviations from the LOOP are highly volatile (Isard 1977; Knetter 1989; Engel and Rogers 1996).

Let $z_{i,t}$ denote deviation from the LOOP for an individual good i:

$$z_{i,t} = f_{i,t} - s_t, \tag{2}$$

where $f_{i,t} \equiv p_{i,t} - p^*_{i,t}$ is the logarithmic difference between the U.S. price and the foreign price of an individual good i and s_t is the logarithmic of the nominal exchange rate. Our empirical analysis centres on the following simple forecasting regression over a k-period horizon:

$$s_{t+k} - s_t = \alpha_{i,k} + \beta_{i,k} z_{i,t} + u^i_{t,t+k}. \tag{3}$$

This is a typical forecast equation used in the international finance literature, with z_t representing the deviation of the log nominal exchange rate from its fundamental value based on a variety of theoretical models; for example, the interest rate differential and monetary fundamentals. In the finance literature, there are also a number of studies examining the long-run predictability of foreign exchange returns using certain instruments that are taken to be proxies for underlying risk factors (e.g., Bekaert and Hodrick 1992; Bauer 2001).

In the sticky-price framework, deviation from the LOOP can be understood as a measure of the price misalignment of an individual good across countries. For instance, an overvalued U.S. dollar (i.e., a decrease in s_t) temporarily causes the price of the good in the United States to be more expensive than in the foreign country (i.e., an increase in $z_{i,t}$). When the U.S. dollar has a tendency to depreciate, such misalignment might be useful in predicting the depreciation of the U.S. dollar (i.e., an increase in s_{t+k} over a k-period horizon). It follows that the slope coefficient in equation (3) is expected to be positive.

The LOOP is the fundamental building block of the PPP condition. There has been extensive research on PPP in the literature suggesting that long-run PPP holds in the post-1973 period for the United States. An important way of examining the empirical content of exchange rate models, however, is to examine their out-of-sample forecasting performance. Evidence of PPP fundamentals beating random walk in out-of-sample forecasting is hardly encouraging (Cheung, Chinn and Pascual 2005). Additionally, monetary models are built upon PPP but assume additional restrictions, and the linkage between exchange rates and monetary fundamentals seems to be tighter than that between exchange rates and PPP fundamentals (Mark and Sul 2001). Nevertheless, it remains true that most studies that claim to have beaten random walk are not robust to refined econometric methods and alternative periods.

In this paper, we examine instead the out-of-sample forecasting power of deviations from the LOOP and show that the superior forecasting performance of these price-misalignment models can be consistent with the poor out-of-sample performance of PPP models. The argument for PPP fundamentals to predict future movements in the exchange rate is based upon the implicit assumption that all relative prices of goods converge to parity at the same speed. But there is little theoretical justification for this assumption. Moreover, empirical evidence suggests that there are a lot of heterogeneous dynamics of deviations from the LOOP for different goods.

With our data sets for U.S.–Japan and U.S.–U.K., we construct deviations from the LOOP for each good in our sample and compute the contemporaneous correlations of each $z_{i,t}$ with $z_{j,t}, j \neq i$ and z_t

(deviations from PPP). Table 1 lists the mean, median and standard deviations of these correlations for each good i. The key message from these statistics is that the price misalignments for individual goods are not highly correlated with each other, and as a result are not highly correlated with the deviation from aggregate PPP. In particular, some pairwise correlations are even negative rather than positive. Therefore, some price misalignments for individual goods may have superior out-of-sample predictive power for future movements in the exchange rate, even though the aggregate PPP fundamental does not.

3 The Data

We use price data obtained from the U.S. Bureau of Labor Statistics, the Japan Statistics Bureau, and the U.K. Office for National Statistics. The data correspond to monthly observations and cover at most the period 1973M01 to 2009M08 for the U.S.–Japan case, and 1987M01 to 2009M08 for the U.S.–U.K. case. However, many observations are missing in the early part of the period for some goods, so we are looking at unbalanced samples in both cases. We collect the good-level price data where available and remove the seasonality of the raw goods price series. This leaves us with a maximum of 440 and 272 time-series observations, respectively.

United States

The U.S. Bureau of Labor Statistics (BLS) publishes price indexes for major groups of consumer expenditures (food and beverages, housing, apparel, transportation, medical care, recreation, education and communications, and other goods and services). The BLS has classified all expenditure items into more than 200 categories. Indexes for all categories are published at the U.S. city average level.

Japan

The Japan Statistics Bureau collects information on prices in a Retail Price Survey, which is conducted in 167 cities, towns and villages. In general, each item encompasses various specifications in terms of quality, volume, container and other characteristics. Goods and services are classified so that each item encompasses similar products in terms of usage, function, etc., and prices within each item are expected to move parallel with each other for long durations.

United Kingdom

The U.K. Office for National Statistics collects good-level price data in its Retail Prices Index (RPI), which is the most familiar general purpose domestic measure of inflation in the United Kingdom. The data set includes details on all consumer spending on goods and services by members of U.K. households.

We report results based on a detailed matching of the data. In the end, we are left with 67 goods for the U.S.–Japan case and 48 goods for the U.S.–U.K. case. For a complete list of the goods, please see the appendix. For both country pairs, the categories consist of highly tradable goods (e.g., women's apparel), goods commonly regarded as non-tradable (e.g., motor vehicle insurance), and goods for which there is wide variation in product differentiation (e.g., spices, seasonings, condiments, sauces). Our sample thus constitutes an interesting cross-section variation, which is key to our analysis since it allows us to identify the heterogeneity in relative price dynamics. Finally, the monthly U.S.-dollar exchange rates per Japanese yen and per U.K. pound are obtained from the International Monetary Fund's International Financial Statistics Database.

4 Empirical Methodology

4.1 Bootstrap procedure

For our analysis, we rely on the bootstrap method to get the p-values of in-sample and out-of-sample statistics of interest. Bootstrap analysis mitigates severe size distortions that may arise from small-sample bias in the estimates of regression coefficients and asymptotic standard errors. In our benchmark case, the data-generating process (DGP) under the null hypothesis that the exchange rate is unpredictable is as follows:

$$\begin{aligned} \Delta s_t &= c_s + \varepsilon_{s,t} \\ z_{i,t} &= c_{i,z} + \phi_{i,1} z_{i.t-1} + ... + \phi_{i,p} z_{i.t-p} + \varepsilon^i_{z,t}, \end{aligned} \tag{4}$$

where $z_{i,t}$ follows a stationary $AR(p)$ process. The lag order of $z_{i,t}$'s process under the null is selected using the Akaike information criterion, given an upper bound of 12 lags. Specifically, the bootstrap algorithm consists of the following five steps.

(i) Construct an empirical probability distribution, which is the non-parametric maximum-likelihood estimate of the population distribution.

(ii) From the empirical distribution function, draw a random sample of size n with replacement of the fitted residuals. The innovation terms are assumed to be i.i.d.

(iii) Based on the DGP model, generate a sequence of pseudo-observations of the same length as the original data series.

(iv) Estimate the regression and calculate the statistic of interest.

(v) Repeat steps (ii) to (iv) 2,000 times, and use the empirical distribution of the 2,000 replications to determine the p-value of the test statistic.

In our bootstrap algorithm, two changes are made to Mark's (1995) bootstrap method, correcting for inconsistencies in the test procedure and for small-sample bias (Kilian 1999). First, when the equation for $z_{i,t}$ is estimated, the small-sample bias correction is taken into account using Shaman and Stine (1988). Second, in the case of the out-of-sample analysis against the random walk model without drift, we restrict the estimate of the drift term in the equation for s_t (i.e., c_s) to zero in generating a sequence of pseudo-observations. In addition, as a robustness check, we conduct bootstrap analysis under the restricted VECM of $z_{i,t}$ and s_t as the null DGP, as suggested by Kilian (1999), such that under the null hypothesis of no exchange rate predictability, the bootstrap DGP is obtained by fitting the restricted VECM:

$$
\begin{aligned}
\Delta s_t &= c_s + \varepsilon_{s,t} \\
\Delta f_{i,t} &= c_{i,z} - h_{i,z} z_{i,t-1} + \sum_{j=1}^{p-1} \phi_{i,1} \Delta s_{t-j} + \sum_{j=1}^{p-1} \phi_{i,2} \Delta f_{i,t-j} + \varepsilon^{i}_{f,t}.
\end{aligned}
\tag{5}
$$

4.2 Testing for superior predictive ability

Since we are simultaneously testing multiple out-of-sample hypotheses in terms of various good prices, the inference based on conventional p-values is likely to be contaminated. As a result of an extensive specification search, data mining is likely to take place. To increase the reliability of our results from the out-of-sample regression, we perform the test of superior predictive ability proposed by Hansen (2005). We have 67 models for the U.S. dollar–Japanese yen exchange rate and 48 models for the U.S. dollar–U.K. pound exchange rate from good-level price misalignments. The SPA test allows us to compare the out-of-sample performance of one benchmark model (the random walk model) to that of a set of alternatives. The SPA test examines the composite null hypothesis that the benchmark model is not inferior to any of the alternatives against the hypothesis that at least one of the linear economic models has superior predictive ability. Empirically, the SPA test consists of the following three steps.

(i) For both the benchmark model and the alternative set of price-misalignment models, forecasts are produced for an evaluation period, $t = 1, ..., N$.

(ii) Let $L(Y_t; \hat{Y}_t)$ denote the loss if one had made the prediction as $\hat{Y}_t$, when the realized value turned out to be Y_t. With $h = 0$ denoting the benchmark forecast, all $h(h = 1, ..., m)$ alternative forecasts are compared with the benchmark via the time series of loss differentials defined as: $X_{h,t} = L(Y_t; \hat{Y}_{0,t}) - L(Y_t; \hat{Y}_{h,t}), h = 1, ..., m, t = 1, ..., N$.

(iii) A test of whether the benchmark model is outperformed by any other model is conducted by testing $H_0 : E[X_h] \leq 0$ for all $h = 1, ..., m$ against $H_A : E[X_h] > 0$ for at least one $h = 1, ..., m$.

In short, a large value for the SPA test statistic represents evidence against the null hypothesis and indicates that at least one model in the model set significantly outperforms the benchmark model.

Therefore, rejecting the null would indicate that at least one price-misalignment model is strictly superior to the random walk. Crucially, this test procedure caters explicitly to the multiple models included in the comparison. Hence, the results are not subject to the criticism of data mining, whereby a sequence of pairwise comparisons between a benchmark model and any set of comparators has a high probability of leading to incorrect rejection of a true null hypothesis.

5 Empirical Results

In this section, we report the empirical results of in-sample and out-of-sample analysis, as well as the results of SPA tests. In-sample analysis gives us some sense of whether *ex post* price misalignments are essential indicators of exchange rate movements. With out-of-sample analysis, we can study whether there is evidence that they are in fact indicators with *ex ante* predictive power.

5.1 Regression estimates and in-sample tests of predictability

Table 2 reports the results of in-sample regressions over six forecast horizons: 3, 6, 12, 24, 36 and 48 months for the U.S. dollar–Japanese yen exchange rate. Table 3 reports the same set of results for the U.S. dollar–U.K. pound exchange rate.[1] At each forecast horizon, we report the R-sq and the estimate of the slope coefficient β of the in-sample regressions in the tables. The t-statistic is then computed based on Newey-West's (1987) HAC covariance matrix estimator with Andrews's (1991) procedure for selecting a truncation lag. Finally, the p-values of t-statistics from bootstrap distributions are plotted in Figures 1 and 2 for the exchange rate of the U.S. dollar vis-à-vis the Japanese yen and the U.K. pound, respectively.

Two overall results are apparent from the in-sample analysis. First, the estimate of the slope coefficient is positive over all horizons for almost all goods in both cases. There are a few exceptions. For example, in the U.S.–Japan case, the estimated β for Good 63 (Motor vehicle insurance) at the 48-month horizon is -0.02. However, it is not significantly different from zero. Figure 3 shows the slope coefficient estimate over the forecast horizon for each good as well as the R-square of the in-sample regression.[2] Both the estimate of β and the R-square tend to increase with the forecast horizon for most goods. Second, over longer horizons, more than 50 per cent of goods display statistical significance in explaining movements in the exchange rate. Specifically, out of 67 goods considered for the U.S. dollar–Japanese yen exchange rate, there are 26, 28, 25, 37, 47 and 37 goods at the 3-, 6-, 12-, 24-, 36- and 48-month forecast horizons, respectively, for which the estimates of their slope coefficients are statistically significant at the 10 per cent level. For the U.S. dollar–U.K. pound exchange rate, there are 22, 19, 17, 26, 13 and 23 goods out of 48 goods at these forecast horizons, for which the estimates of the slope coefficients are statistically significant at the 10 per cent level.

[1] The result for the 1-month horizon is not reported to save space, but is available upon request from the authors.

[2] Goods are ordered in the magnitude of the estimate of their slope coefficients at the 24-month horizon.

5.2 Out-of-sample tests of predictability

Many price misalignments for individual goods seem to be significant indicators of exchange rate movements. Next, we perform out-of-sample analysis to examine whether there is evidence of exchange rate predictability. Tables 4 and 5, respectively, report the results for the U.S. dollar–Japanese yen exchange rate and the U.S. dollar–U.K. pound exchange rate from the out-of-sample regression versus (i) the random walk (RW) with no drift and (ii) the RW with drift. The tables show the t-statistics from Clark and West's (2007) procedure for testing for the equal predictive ability of two nested models CW(A).[3] The p-values from bootstrap distributions are plotted in Figures 4 and 5 for various forecast horizons.

U.S. dollar–Japanese yen exchange rate

For the out-of-sample analysis, the date at which the first forecast is made is generally selected at 1983M01. For goods where observations are available after, or just several years before, 1983M01, however, the date for the first forecast is chosen as the midpoint of their available sample period. Our results indicate that (i) there are 16, 8, 5, 18, 22 and 10 goods over the 3-, 6-, 12-, 24-, 36- and 48-month forecast horizons, respectively, that perform better than the RW with no drift at the 10 per cent level of significance, and (ii) there are 12, 13, 6, 23, 39 and 29 goods over the same six horizons, respectively, that perform better than the RW with drift. Generally, over longer horizons (24 and 36 months), there is stronger evidence of exchange rate predictability. But the trend is non-linear; as at the 48-month forecast horizon, a lot of statistical significance is lost.[4]

When comparing the out-of-sample regression results between the RW with no drift and the RW with drift, it is clear that the price-misalignment models for individual goods beat the RW with drift more frequently than the RW with no drift, particularly over long horizons. This is because the model for random walk with no drift is a better representation of changes in the bilateral exchange rate between the U.S. dollar and the Japanese yen over longer forecast horizons. We compute the ratios of the root-mean-square prediction error (RMSPE) for the driftless RW model to the RMSPE for the RW with drift model and report them in Table 6.[5] The RMSPE ratios are generally smaller than 1 over longer horizons.

U.S. dollar–U.K. pound exchange rate

We perform similar analysis for the bilateral exchange rate between the U.S. dollar and the U.K. pound. However, the results are not as positive. In particular, out of 48 goods, there are only 3, 1, 1, 5,

[3] CW(A) represents the t-statistic based on Newey-West's (1987) HAC covariance matrix estimator with Andrews's (1991) procedure for selecting a truncation lag, so as to account for serial correlation when the forecast horizon is more than one period. The results are robust to the standard Clark and West t-statistic.

[4] This could be related to the length of the data.

[5] This RMSPE ratio depends on the forecast period (i.e., the date for the first forecast) as well as the starting date of observations.

0 and 2 goods over the 3-, 6-, 12-, 24-, 36- and 48-month forecast horizons, respectively, that perform better than the RW with no drift at the 10 per cent level of significance; on the other hand, there are 3, 2, 1, 6, 0 and 2 goods over these forecast horizons that perform better than the RW with drift at the 10 per cent level of significance. The results suggest that from a real-time forecaster's point of view, price-misalignment models for individual goods may not be as useful in predicting the bilateral exchange rate between the U.S. dollar and the U.K. pound as in predicting the exchange rate between the U.S. dollar and the Japanese yen. However, we should not discount the in-sample fit results. With in-sample analysis, we use the full sample in fitting the models of interest. With out-of-sample analysis, we mimic data constraints faced by real-time forecasters. In practice, in-sample tests tend to reject the null hypothesis of no predictability more often than out-of-sample tests. But that is not necessarily an indication that in-sample tests are biased in favour of detecting spurious predictability (Inoue and Kilian 2004). Rather, out-of-sample analysis based on sample splitting involves loss of information and therefore, perhaps, lower power in small samples like the one we have for the U.S.–U.K. case.

5.3 SPA tests

To account for potential data mining, we conduct SPA tests for the U.S.–Japan case by grouping 67 goods upon availability of their observations. The SPA test requires that both a set of alternative models and the benchmark model have the same number of forecasts. Since we have an unbalanced sample due to data availability, we perform the SPA test for six cases as listed in Table 7. For instance, 25 goods that have observations available from 1973M03 are considered as one group, so that the result of the SPA test for such a group can be consistently complementary in understanding the out-of-sample regression results for those 25 goods.[6]

Tables 8 and 9 report the results of the SPA test for the six cases considered. Together with the consistent p-values, we also report the upper and lower bounds, as well the critical values at the 10 per cent, 5 per cent and 1 per cent level.[7] The upper bound is the p-value of a conservative test which assumes that all the competing models are precisely as good as the benchmark in terms of the expected loss. The lower bound is the p-value of a liberal test whose null hypothesis assumes that the models with worse performance than the benchmark are poor models in the limit. Therefore, these can be viewed as asymptotic upper and lower bounds for the actual p-value.

In case 1, where all 25 goods have data available back to 1973M03, SPA test results suggest that our price-misalignment models beat driftless RW at the 6-month horizon and longer, and beat RW with drift at the 12-month horizon and longer, at a 5 per cent level of significance. For the 13 goods

[6] When goods are mixed in terms of availability of observations, and the sample used for the out-of-sample regressions of goods is different, there is an issue in choosing the starting date of the sample period of the random walk with drift as a benchmark model in the SPA test. In our benchmark case, such a starting date is set equal to the starting date of the longest sample of a good. For instance, in case 4, the sample for the RW with drift starts from 1973M03. Our results are robust to alternative choices of dates.

[7] P-values are designed to control for the size of a test and are not informative of the power of a test. When one fails to reject the null hypothesis, the critical value can be informative about the power of a test. Critical values that are large indicate that the data being analyzed are not very informative about the hypothesis of interest, and that the SPA test may lack power.

in case 2, price-misalignment models beat both RWs at the 6-month horizon at a 5 per cent level of significance. In both cases 1 and 2, it seems to be easier for the deviation from the LOOP models to beat RW without drift than with drift. In case 3, however, where all 26 goods have observations starting only from 1997M12, the opposite is true. Still, our benchmark models beat the null RWs at the 12-month horizon or higher. Next, in cases 4, 5 and 6, where we mix groups together and employ various lengths of the samples, it follows that we can reject the null (RW) of the SPA test at the 5 per cent level of significance over the 12-month horizons for all cases. In case 6, where we restrict the sample to start only from 1997M12, the price-misalignment models even beat both RWs at all horizons at the 5 per cent level. The SPA test results indicate that at least one of our price-misalignment models has superior predictive ability over the RW models both with and without drift, after accounting for potential data snooping.

5.4 Goods whose price misalignments can predict exchange rate changes

We see that, upon bringing the price-misalignment model to a disaggregated level, there is a lot of heterogeneity in terms of predictive power for future exchange rates. The next question is, then, which particular good-level price misalignment can predict changes in the exchange rate? For the U.S.–Japan case, we examine 67 good-level price-misalignment models in terms of their predictability both in-sample and out-of-sample. Our data cover a broad selection of consumption goods. Some of them provide good in-sample fit, and others display superior out-of-sample predictability over random walks for forecasting movements in the nominal exchange rate. Looking closely at these goods, several observations can be made.

First, among the items we study for price misalignment, there is a great amount of heterogeneity in terms of price sluggishness. Some prices are very flexible, such as most food items (e.g., pork chops, lettuce). Others are quite sticky, such as intercity bus fare.[8] Moreover, there is also heterogeneity in terms of the persistence of price misalignment for each good. Although we find that most price dispersions for the 67 goods are quite persistent, some exceptions apply. For example, the first-order autoregressive coefficient for lettuce is only 0.69, compared to 0.98 for utility (piped) gas service. Results from our forecasting exercises suggest that the stickiness of prices or the persistence of cross-border price misalignment are not necessarily relevant to whether a good-level price-misalignment model has superior predictive power for changes in the exchange rate.

Second, out of the 67 goods that we have price data for, 14 are actually non-tradable. At first glance, one may think only tradable-good price misalignments across countries might have predictive power for future movements in the exchange rate, since movements in relative prices can trigger expenditure switching effects across borders. But, in fact, our in-sample estimation results show that, among the 14 non-tradable goods included in our sample for the U.S.–Japan case, 12 display both economic and statistical significance in long-horizon (36-month) predictability. Even at a short (3-month) forecast horizon, 8 out of 14 non-tradable goods provide in-sample fit at a 10 per cent level of significance.

[8] For more information on price stickiness for each good, refer to the micro studies on the U.S. data by Bils and Klenow (2004).

For out-of-sample predictability, fewer non-tradable goods display significant forecasting power. But there are always some non-tradable price-misalignment models that beat random walk models at various forecast horizons; for example, price misalignment for electricity, utility (piped) gas service and intracity transportation can all beat RWs at certain forecast horizons, and in some cases, at all forecast horizons.

The price misalignment on some non-tradable items that display predictive power for exchange rates is not surprising. In fact, we should expect non-tradable price dispersions to be better at forecasting future movements in the exchange rate. For tradable goods, a large price difference across borders may signal that subsequent adjustment is taking place either through price changes or nominal exchange rate changes. For non-tradable items, however, the adjustment has to take place through changes in the exchange rate, since there is no trade channel. Thus, for certain goods, currency misalignments may be indicated by price deviations, which could further help to predict future movements in the exchange rate.

To get a clearer picture, we sort out the goods whose price misalignments have significant predictive power for changes in the nominal exchange rate both in-sample and out-of-sample at a 10 per cent level of significance, and provide the list of goods in Table 10 for various forecast horizons. We emphasize two observations. First, electricity can predict the bilateral exchange rate between the U.S. dollar and the Japanese yen both in-sample and out-of-sample at all forecast horizons. Most electricity in both the United States and Japan is generated using coal, oil and natural gas. Global energy prices may provide a natural mean reversion target for the misalignment of electricity prices. Second, frozen fish and seafood also display significant predictive power for the nominal exchange rate at the 3-, 6-, 12-, 24- and 36-month forecast horizons. Japan is the top export market for U.S. fish and seafood, accounting for about a quarter of its total exports. As suggested by Crucini, Telmer and Zachariadis (2005), the tradability of a good may be negatively related to the good-by-good measures of cross-sectional price dispersion. All retail goods involve significant amounts of non-traded inputs. However, the more tradable a good is, the more impact arbitrage conditions have on its relative price across borders.

6 Conclusion

In this paper, we examine out-of-sample exchange rate predictability based on price misalignment at the good level. We find that, for many goods, our benchmark model outperforms random walks either with or without drift at the 5 per cent level of significance over long horizons. When we apply tests of superior predictive ability, taking into account small-sample bias and data-mining issues, we find that, starting from the 12-month horizon, the price-misalignment model outperforms both random walks at the 5 per cent level of significance. Our results are robust to alternative sets of goods and sample periods.

Our findings have potentially important implications. First, our model generates robust out-of-sample exchange rate predictability. Bringing in insights from the micro-level data, our findings suggest that price misalignment for some goods (for example, electricity, frozen fish and seafood) even has predictive power for exchange rates at very short horizons (3 months), which is of practical relevance over horizons of concern to policy-makers. Second, we highlight the importance of heterogeneity at the micro level for understanding the macroeconomy. Good-level relative prices not only are impacted by exchange rate fluctuations, but also show predictive power for their future values when heterogeneity is accounted for. Finally, our forecasting exercises certainly do not provide conclusive evidence that price-misalignment models determine the exchange rate; rather, it is the currency misalignment itself, through price dispersions, that helps to predict the exchange rate's subsequent adjustment. In addition, there may well be room for monetary models (Engel, Mark and West 2007), models with heterogeneous information (Bacchetta and van Wincoop 2006), or models based on the microstructure of foreign exchange markets (Evans and Lyons 2002) to improve our understanding of currency movements.

References

Andrews, D. 1991. “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation.” *Econometrica* 59 (3): 817–58.

Bacchetta, P. and E. van Wincoop. 2006. “Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?” *American Economic Review* 96 (3): 552–76.

Bauer, G. 2001. “The Foreign Exchange Risk Premium Over the Long Run.” University of Rochester Working Paper.

Bekaert, G. and R. Hodrick. 1992. “Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets.” *Journal of Finance* 47 (2): 467–509.

Berkowitz, J. and L. Giorgianni. 2001. “Long-Horizon Exchange Rate Predictability?” *Review of Economics and Statistics* 83 (1): 81–91.

Bils, M. and P. Klenow. 2004. “Some Evidence on the Importance of Sticky Prices.” *Journal of Political Economy* 112 (5): 947–85.

Cheung, Y.-W., M. Chinn and A. Pascual. 2005. “Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?” *Journal of International Money and Finance* 24 (7): 1150–75.

Clark, T. and M. McCracken. 2001. “Tests of Equal Forecast Accuracy and Encompassing for Nested Models.” *Journal of Econometrics* 105 (1): 85–110.

Clark, T. and K. West. 2007. “Approximately Normal Tests for Equal Predictive Accuracy in Nested Models.” *Journal of Econometrics* 138 (1): 291–311.

Corradi, V. and N. Swanson. 2004. “Some Recent Developments in Predictive Accuracy Testing with Nested Models and (Generic) Nonlinear Alternatives.” *International Journal of Forecasting* 20 (2): 185–99.

——. 2007. “Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes.” *International Economic Review* 48 (1): 67–109.

Crucini, M., C. Telmer and M. Zachariadis. 2005. “Understanding European Real Exchange Rates.” *American Economic Review* 95 (3): 724–38.

Diebold, F. and R. Mariano. 1995. “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics* 13 (3): 253–63.

Engel, C., N. Mark and K. West. 2007. “Exchange Rate Models Are Not as Bad as You Think.” *NBER Macroeconomics Annual* edited by D. Acemoglu, K. Rogoff and M. Woodford, 381–441. Chicago: University of Chicago Press.

Engel, C. and J. Rogers. 1996. “How Wide Is the Border?” *American Economic Review* 86 (5):

1112–25.

Evans, M. and R. Lyons. 2002. “Order Flow and Exchange Rate Dynamics.” *Journal of Political Economy* 110 (1): 170–80.

Gourinchas, P.-O. and H. Rey. 2007. “International Financial Adjustment.” *Journal of Political Economy* 115 (4): 665–703.

Hansen, P. 2005. “A Test for Superior Predictive Ability.” *Journal of Business and Economic Statistics* 23 (4): 365–80.

Imbs, J., H. Mumtaz, M. Ravn and H. Rey. 2005. “PPP Strikes Back: Aggregation and the Real Exchange Rate.” *Quarterly Journal of Economics* 120 (1): 1–43.

Inoue, A. and L. Kilian. 2004. “In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?” *Econometric Reviews* 23 (4): 371–402.

Isard, P. 1977. “How Far Can We Push the ‘Law of One Price’?” *American Economic Review* 67 (5): 942–48.

Kilian, L. 1999. “Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?” *Journal of Applied Econometrics* 14 (5): 491–510.

Knetter, M. 1989. “Price Discrimination by U.S. and German Exporters.” *American Economic Review* 79 (1): 198–210.

Mark, N. 1995. “Exchange Rates and Fundamentals: Evidence on Long Horizon Predictability.” *American Economic Review* 85 (1): 201–18.

Mark, N. and D. Sul. 2001. “Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel.” *Journal of International Economics* 53 (1): 29–52.

Meese, R. and K. Rogoff. 1983. “Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?” *Journal of International Economics* 14 (1-2): 3–24.

Molodtsova, T. and D. Papell. 2009. “Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals.” *Journal of International Economics* 77 (2): 167–80.

Newey, W. and K. West. 1987. “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55 (3): 703–08.

Rogoff, K. and V. Stavrakeva. 2008. “The Continuing Puzzle of Short Horizon Exchange Rate Forecasting.” NBER Working Paper No. 14071.

Shaman, P. and R. Stine. 1988. “The Bias of Autoregressive Coefficient Estimators.” *Journal of the American Statistical Association* 83 (403): 842–48.

White, H. 2000. “A Reality Check for Data Snooping.” *Econometrica* 68 (5): 1097–1126.

Table 1: Correlations between Deviations from the LOOP

US–Japan				US–UK			
Goods	Mean	Median	Std	Goods	Mean	Median	Std
Aggregate	0.78	0.88	0.26	Aggregate	0.63	0.72	0.29
Good 01	0.72	0.85	0.30	Good 01	0.29	0.35	0.44
Good 02	0.61	0.67	0.29	Good 02	0.50	0.57	0.34
Good 03	0.66	0.72	0.21	Good 03	0.65	0.72	0.28
Good 04	0.52	0.52	0.23	Good 04	0.52	0.70	0.43
Good 05	0.74	0.79	0.17	Good 05	0.70	0.81	0.28
Good 06	0.63	0.70	0.23	Good 06	0.64	0.72	0.27
Good 07	0.66	0.80	0.34	Good 07	0.53	0.56	0.25
Good 08	0.63	0.70	0.29	Good 08	0.35	0.35	0.22
Good 09	0.71	0.78	0.24	Good 09	0.59	0.57	0.23
Good 10	0.62	0.63	0.19	Good 10	0.37	0.43	0.23
Good 11	0.68	0.71	0.20	Good 11	0.56	0.73	0.41
Good 12	0.75	0.80	0.19	Good 12	0.69	0.71	0.20
Good 13	0.65	0.81	0.37	Good 13	0.60	0.69	0.28
Good 14	0.52	0.59	0.20	Good 14	0.41	0.62	0.44
Good 15	0.71	0.83	0.30	Good 15	0.29	0.34	0.33
Good 16	0.67	0.73	0.19	Good 16	0.45	0.64	0.44
Good 17	0.66	0.74	0.29	Good 17	0.63	0.70	0.27
Good 18	0.73	0.84	0.28	Good 18	0.57	0.66	0.25
Good 19	0.49	0.53	0.14	Good 19	0.64	0.65	0.17
Good 20	0.68	0.82	0.34	Good 20	0.42	0.54	0.47
Good 21	0.53	0.61	0.22	Good 21	0.70	0.83	0.33
Good 22	0.32	0.36	0.18	Good 22	0.55	0.59	0.28
Good 23	0.62	0.79	0.40	Good 23	0.47	0.56	0.41
Good 24	0.62	0.77	0.37	Good 24	0.40	0.43	0.24
Good 25	0.55	0.68	0.36	Good 25	0.53	0.56	0.22
Good 26	0.69	0.79	0.26	Good 26	0.46	0.54	0.42
Good 27	0.70	0.80	0.33	Good 27	0.36	0.40	0.32
Good 28	0.64	0.73	0.27	Good 28	0.74	0.85	0.27
Good 29	0.74	0.85	0.27	Good 29	0.61	0.76	0.40
Good 30	0.75	0.86	0.23	Good 30	0.75	0.82	0.23
Good 31	0.75	0.84	0.24	Good 31	0.66	0.83	0.37
Good 32	0.73	0.82	0.22	Good 32	0.35	0.41	0.45
Good 33	0.75	0.88	0.31	Good 33	0.62	0.76	0.32
Good 34	0.73	0.79	0.22	Good 34	0.57	0.67	0.33
Good 35	0.60	0.78	0.42	Good 35	0.63	0.67	0.24
Good 36	0.65	0.74	0.32	Good 36	0.48	0.56	0.36
Good 37	0.70	0.82	0.32	Good 37	-0.11	-0.14	0.28
Good 38	0.61	0.70	0.26	Good 38	0.43	0.54	0.48
Good 39	0.52	0.62	0.29	Good 39	0.37	0.48	0.44
Good 40	0.74	0.84	0.27	Good 40	0.44	0.53	0.46
Good 41	0.69	0.78	0.30	Good 41	0.32	0.39	0.54
Good 42	0.65	0.74	0.30	Good 42	0.41	0.49	0.40
Good 43	0.67	0.74	0.24	Good 43	0.07	0.09	0.53
Good 44	0.62	0.77	0.37	Good 44	0.61	0.64	0.21
Good 45	0.69	0.81	0.34	Good 45	0.19	0.22	0.52
Good 46	0.10	-0.03	0.45	Good 46	0.68	0.83	0.36
Good 47	0.46	0.67	0.47	Good 47	0.67	0.79	0.32
Good 48	0.16	0.16	0.51	Good 48	0.68	0.69	0.20
Good 49	-0.01	-0.16	0.47				
Good 50	0.33	0.21	0.36				
Good 51	0.69	0.71	0.18				
Good 52	0.39	0.41	0.40				
Good 53	0.43	0.41	0.41				
Good 54	0.61	0.66	0.30				
Good 55	0.58	0.63	0.30				
Good 56	0.44	0.43	0.38				
Good 57	0.52	0.52	0.31				
Good 58	0.74	0.85	0.26				
Good 59	0.51	0.62	0.37				
Good 60	0.56	0.68	0.31				
Good 61	0.69	0.78	0.25				
Good 62	0.72	0.88	0.34				
Good 63	0.21	0.28	0.46				
Good 64	0.64	0.77	0.32				
Good 65	0.62	0.80	0.41				
Good 66	0.74	0.84	0.27				
Good 67	0.27	0.32	0.32				

Table 2: Results of In-sample Regression (US – Japan)

		3-month		6-month		12-month		24-month		36-month		48-month	
Goods	Starting Date	R-sq	Beta	R-sq	Beta	R-sq	Beta	R-sq	Beta	R-sq	Beta	R-sq	Beta
Good 01	1973.03	0.00	0.01	0.01	0.03	0.02	0.07	0.08	0.20	0.18	0.34	0.30	0.46
Good 02	1973.03	0.01	0.03	0.03	0.07	0.05	0.14	0.16	0.37	0.30	0.60	0.43	0.76
Good 03	1973.03	0.01	0.03	0.02	0.07	0.04	0.16	0.17	0.47	0.32	0.75	0.41	0.90
Good 04	1997.12	0.02	0.09	0.03	0.16	0.06	0.30	0.52	1.32	0.59	1.16	0.33	0.48
Good 05	1980.01	0.03	0.06	0.05	0.12	0.11	0.23	0.25	0.53	0.43	0.81	0.54	0.95
Good 06	1977.12	0.04	0.09	0.08	0.18	0.17	0.35	0.27	0.63	0.37	0.89	0.44	1.07
Good 07	1997.12	0.06	0.14	0.10	0.24	0.15	0.37	0.47	0.94	0.74	1.18	0.48	0.64
Good 08	1977.12	0.01	0.02	0.02	0.04	0.03	0.07	0.08	0.17	0.16	0.28	0.25	0.37
Good 09	1973.03	0.01	0.03	0.03	0.07	0.06	0.15	0.15	0.33	0.28	0.52	0.40	0.67
Good 10	1973.03	0.01	0.02	0.02	0.04	0.04	0.09	0.11	0.20	0.18	0.29	0.27	0.39
Good 11	1977.12	0.01	0.03	0.03	0.07	0.07	0.15	0.15	0.32	0.24	0.49	0.33	0.61
Good 12	1977.12	0.02	0.05	0.04	0.10	0.09	0.20	0.20	0.42	0.28	0.60	0.35	0.71
Good 13	1997.12	0.07	0.12	0.12	0.22	0.19	0.35	0.43	0.80	0.60	0.92	0.40	0.48
Good 14	1973.03	0.00	0.02	0.02	0.06	0.04	0.13	0.14	0.34	0.17	0.44	0.26	0.57
Good 15	1973.03	0.00	0.02	0.01	0.04	0.03	0.09	0.10	0.24	0.18	0.41	0.25	0.50
Good 16	1997.12	0.08	0.14	0.19	0.28	0.46	0.54	0.78	1.00	0.49	0.71	0.21	0.26
Good 17	1977.12	0.01	0.03	0.02	0.06	0.04	0.11	0.13	0.29	0.22	0.46	0.29	0.56
Good 18	1977.12	0.02	0.04	0.04	0.08	0.07	0.14	0.16	0.29	0.24	0.43	0.29	0.51
Good 19	1973.03	0.00	0.02	0.00	0.03	0.01	0.06	0.04	0.25	0.08	0.42	0.14	0.56
Good 20	1997.12	0.08	0.15	0.16	0.29	0.30	0.49	0.65	1.06	0.73	1.11	0.35	0.49
Good 21	1973.03	0.00	0.01	0.01	0.03	0.02	0.06	0.05	0.14	0.09	0.20	0.14	0.28
Good 22	1973.03	0.00	0.00	0.00	0.00	0.01	0.03	0.02	0.07	0.03	0.09	0.05	0.13
Good 23	1997.12	0.06	0.10	0.11	0.19	0.20	0.32	0.44	0.70	0.52	0.73	0.31	0.36
Good 24	1997.12	0.10	0.16	0.17	0.29	0.31	0.49	0.64	1.03	0.69	0.99	0.31	0.42
Good 25	1973.03	0.01	0.04	0.03	0.08	0.06	0.15	0.10	0.29	0.16	0.44	0.19	0.53
Good 26	1973.03	0.01	0.03	0.02	0.06	0.05	0.13	0.11	0.28	0.20	0.45	0.31	0.60
Good 27	1997.12	0.07	0.15	0.14	0.28	0.25	0.46	0.53	0.95	0.56	0.94	0.32	0.45
Good 28	1973.03	0.02	0.04	0.05	0.10	0.09	0.19	0.18	0.37	0.21	0.47	0.29	0.58
Good 29	1997.12	0.09	0.17	0.16	0.32	0.29	0.53	0.65	1.10	0.68	1.04	0.33	0.44
Good 30	1978.01	0.02	0.05	0.06	0.12	0.12	0.24	0.23	0.48	0.34	0.71	0.43	0.86
Good 31	1997.12	0.09	0.19	0.17	0.36	0.29	0.58	0.67	1.22	0.73	1.18	0.36	0.50
Good 32	1997.12	0.14	0.25	0.26	0.47	0.45	0.77	0.79	1.43	0.66	1.23	0.26	0.47
Good 33	1973.03	0.02	0.03	0.04	0.07	0.08	0.15	0.20	0.33	0.32	0.48	0.42	0.59
Good 34	1997.12	0.11	0.21	0.18	0.38	0.27	0.60	0.57	1.18	0.54	1.02	0.30	0.44
Good 35	1997.12	0.04	0.07	0.06	0.12	0.10	0.19	0.27	0.46	0.40	0.53	0.30	0.31
Good 36	1973.03	0.03	0.06	0.06	0.13	0.15	0.29	0.34	0.63	0.48	0.93	0.53	1.08
Good 37	1973.03	0.01	0.03	0.03	0.06	0.05	0.13	0.13	0.28	0.24	0.45	0.31	0.55
Good 38	1973.03	0.02	0.03	0.05	0.08	0.09	0.16	0.21	0.35	0.34	0.53	0.47	0.71
Good 39	1973.03	0.00	0.01	0.01	0.04	0.02	0.09	0.08	0.27	0.18	0.49	0.29	0.68
Good 40	1997.12	0.10	0.19	0.17	0.33	0.28	0.53	0.64	1.13	0.72	1.13	0.35	0.48
Good 41	1997.12	0.04	0.10	0.08	0.19	0.13	0.30	0.41	0.75	0.59	0.79	0.34	0.37
Good 42	1997.12	0.02	0.06	0.04	0.11	0.10	0.23	0.47	0.71	0.77	0.83	0.43	0.39
Good 43	1977.12	0.03	0.05	0.06	0.10	0.10	0.18	0.22	0.38	0.33	0.55	0.40	0.66
Good 44	1997.12	0.06	0.10	0.11	0.19	0.20	0.32	0.49	0.74	0.66	0.92	0.53	0.62
Good 45	1997.12	0.05	0.10	0.09	0.18	0.18	0.30	0.46	0.68	0.57	0.74	0.33	0.36
Good 46	1997.12	0.01	0.02	0.02	0.07	0.04	0.14	0.21	0.56	0.22	0.50	0.08	0.17
Good 47	1997.12	0.02	0.03	0.04	0.05	0.08	0.10	0.19	0.24	0.27	0.28	0.19	0.17
Good 48	1977.12	0.01	0.02	0.02	0.04	0.04	0.07	0.13	0.20	0.27	0.36	0.40	0.49
Good 49	1997.12	0.01	0.02	0.02	0.06	0.03	0.10	0.09	0.28	0.06	0.21	0.02	0.08
Good 50	1997.12	0.05	0.13	0.12	0.28	0.25	0.56	0.55	1.19	0.37	0.78	0.14	0.27
Good 51	1997.12	0.06	0.16	0.14	0.33	0.31	0.66	0.80	1.46	0.79	1.28	0.37	0.52
Good 52	1977.12	0.01	0.01	0.02	0.03	0.01	0.04	0.05	0.09	0.11	0.17	0.16	0.22
Good 53	1977.12	0.01	0.03	0.03	0.06	0.04	0.09	0.11	0.22	0.20	0.36	0.27	0.45
Good 54	1973.03	0.02	0.03	0.04	0.06	0.07	0.12	0.19	0.28	0.33	0.43	0.43	0.51
Good 55	1977.12	0.01	0.02	0.02	0.05	0.03	0.07	0.09	0.18	0.17	0.28	0.23	0.35
Good 56	1977.12	0.01	0.02	0.03	0.05	0.04	0.08	0.12	0.20	0.21	0.32	0.29	0.41
Good 57	1977.12	0.01	0.02	0.03	0.05	0.05	0.09	0.12	0.20	0.23	0.32	0.32	0.41
Good 58	1997.12	0.10	0.20	0.18	0.36	0.27	0.56	0.61	1.16	0.66	1.11	0.35	0.49
Good 59	1973.03	0.02	0.04	0.04	0.09	0.08	0.19	0.21	0.45	0.35	0.67	0.44	0.80
Good 60	1973.03	0.00	0.01	0.01	0.03	0.01	0.05	0.06	0.19	0.13	0.35	0.20	0.48
Good 61	1973.03	0.01	0.02	0.02	0.04	0.04	0.09	0.12	0.22	0.21	0.35	0.31	0.44
Good 62	1973.03	0.01	0.02	0.02	0.04	0.04	0.07	0.11	0.17	0.19	0.26	0.29	0.34
Good 63	1973.03	0.00	0.01	0.01	0.02	0.01	0.04	0.01	0.05	0.00	0.03	0.00	-0.02
Good 64	1997.12	0.04	0.09	0.07	0.16	0.12	0.25	0.31	0.57	0.40	0.63	0.41	0.47
Good 65	1997.12	0.05	0.09	0.09	0.17	0.14	0.27	0.37	0.64	0.53	0.74	0.38	0.42
Good 66	1978.01	0.03	0.05	0.05	0.11	0.12	0.22	0.28	0.49	0.41	0.71	0.49	0.85
Good 67	1973.03	0.01	0.03	0.02	0.06	0.04	0.11	0.07	0.21	0.07	0.25	0.06	0.26
PPP	1973.03	0.02	0.04	0.03	0.08	0.08	0.17	0.20	0.38	0.33	0.58	0.44	0.71

Table 3: Results of In-sample Regression (US – UK)

		3-month		6-month		12-month		24-month		36-month		48-month	
Goods	Starting Date	R-sq	Beta	R-sq	Beta	R-sq	Beta	R-sq	Beta	R-sq	Beta	R-sq	Beta
Good 01	1987.01	0.05	0.09	0.09	0.18	0.10	0.25	0.23	0.44	0.32	0.54	0.41	0.72
Good 02	1987.01	0.06	0.11	0.13	0.23	0.17	0.33	0.16	0.43	0.24	0.60	0.42	0.89
Good 03	1987.01	0.12	0.22	0.22	0.43	0.30	0.67	0.49	1.03	0.56	1.16	0.66	1.38
Good 04	1987.01	0.04	0.07	0.07	0.14	0.12	0.23	0.28	0.41	0.32	0.45	0.41	0.56
Good 05	1997.12	0.09	0.15	0.17	0.31	0.25	0.50	0.46	1.04	0.46	1.27	0.53	1.52
Good 06	1997.12	0.02	0.07	0.05	0.16	0.10	0.33	0.14	0.54	0.03	0.29	0.13	0.67
Good 07	1987.01	0.07	0.14	0.16	0.31	0.22	0.46	0.34	0.71	0.41	0.85	0.49	1.03
Good 08	1987.01	0.05	0.06	0.10	0.13	0.15	0.20	0.11	0.21	0.17	0.27	0.15	0.27
Good 09	1987.01	0.08	0.17	0.16	0.34	0.20	0.50	0.25	0.69	0.32	0.85	0.47	1.14
Good 10	1987.01	0.08	0.13	0.17	0.28	0.30	0.47	0.19	0.45	0.10	0.36	0.15	0.48
Good 11	1987.01	0.04	0.07	0.07	0.13	0.13	0.23	0.27	0.38	0.31	0.43	0.40	0.53
Good 12	1997.12	0.08	0.16	0.15	0.36	0.28	0.78	0.52	1.41	0.51	1.44	0.66	1.74
Good 13	1987.01	0.07	0.13	0.15	0.25	0.25	0.42	0.38	0.65	0.46	0.76	0.59	0.94
Good 14	1987.01	0.01	0.02	0.02	0.05	0.05	0.12	0.12	0.22	0.15	0.26	0.27	0.38
Good 15	1987.01	0.03	0.06	0.09	0.14	0.15	0.26	0.28	0.41	0.16	0.32	0.21	0.43
Good 16	1987.01	0.01	0.03	0.02	0.06	0.06	0.12	0.16	0.22	0.21	0.26	0.32	0.36
Good 17	1987.01	0.06	0.11	0.11	0.22	0.20	0.37	0.36	0.59	0.50	0.73	0.69	0.94
Good 18	1987.01	0.04	0.13	0.07	0.24	0.14	0.45	0.32	0.80	0.24	0.73	0.36	0.99
Good 19	1997.12	0.07	0.13	0.15	0.30	0.27	0.55	0.66	1.15	0.55	1.11	0.48	1.11
Good 20	1987.01	0.05	0.09	0.09	0.16	0.13	0.24	0.09	0.24	0.06	0.22	0.03	0.18
Good 21	1997.12	0.04	0.08	0.08	0.17	0.13	0.30	0.17	0.48	0.13	0.49	0.34	0.91
Good 22	1987.01	0.09	0.12	0.17	0.23	0.32	0.41	0.49	0.65	0.56	0.83	0.63	0.99
Good 23	1987.01	0.07	0.10	0.13	0.20	0.19	0.31	0.21	0.40	0.19	0.44	0.13	0.40
Good 24	1987.01	0.02	0.04	0.06	0.10	0.17	0.22	0.44	0.41	0.42	0.43	0.53	0.55
Good 25	1987.01	0.03	0.12	0.14	0.33	0.18	0.49	0.24	0.70	0.29	0.80	0.32	0.94
Good 26	1987.01	0.05	0.08	0.10	0.16	0.16	0.26	0.14	0.33	0.10	0.34	0.12	0.43
Good 27	1997.12	0.00	0.02	0.01	0.06	0.02	0.20	0.22	0.80	0.26	0.85	0.61	1.38
Good 28	1997.12	0.07	0.12	0.17	0.28	0.25	0.47	0.30	0.73	0.31	0.78	0.60	1.18
Good 29	1997.12	0.02	0.04	0.06	0.09	0.11	0.17	0.08	0.21	0.06	0.21	0.23	0.50
Good 30	1997.12	0.06	0.14	0.14	0.32	0.24	0.60	0.41	1.07	0.38	1.12	0.60	1.52
Good 31	1997.12	0.03	0.05	0.06	0.12	0.12	0.22	0.13	0.33	0.12	0.38	0.34	0.75
Good 32	1987.01	0.02	0.05	0.04	0.10	0.07	0.18	0.24	0.41	0.38	0.53	0.51	0.70
Good 33	1987.01	0.07	0.14	0.12	0.27	0.22	0.48	0.41	0.78	0.46	0.87	0.60	1.09
Good 34	1987.01	0.08	0.14	0.15	0.28	0.25	0.49	0.46	0.78	0.54	0.88	0.70	1.10
Good 35	1987.01	0.07	0.15	0.14	0.30	0.23	0.50	0.34	0.76	0.32	0.82	0.51	1.13
Good 36	1987.01	0.05	0.11	0.10	0.22	0.21	0.47	0.43	0.79	0.50	0.88	0.67	1.14
Good 37	1987.01	0.00	0.01	0.00	0.01	0.01	0.07	0.02	0.12	0.00	0.04	0.00	-0.06
Good 38	1987.01	0.04	0.06	0.08	0.12	0.13	0.21	0.09	0.23	0.06	0.22	0.05	0.26
Good 39	1987.01	0.09	0.14	0.17	0.28	0.22	0.41	0.10	0.35	0.00	0.02	0.01	-0.11
Good 40	1987.01	0.07	0.11	0.13	0.21	0.18	0.32	0.12	0.34	0.05	0.28	0.04	0.29
Good 41	1987.01	0.02	0.04	0.05	0.07	0.07	0.11	0.02	0.08	0.00	0.03	0.00	-0.02
Good 42	1987.01	0.05	0.11	0.08	0.22	0.17	0.46	0.38	0.83	0.47	0.96	0.65	1.27
Good 43	1987.01	0.00	0.00	0.00	0.00	0.00	0.01	0.07	0.08	0.16	0.13	0.27	0.19
Good 44	1997.12	0.08	0.12	0.16	0.29	0.32	0.61	0.64	1.14	0.59	1.13	0.55	1.15
Good 45	1987.01	0.00	0.01	0.01	0.01	0.04	0.03	0.00	0.01	0.01	-0.02	0.03	-0.05
Good 46	1997.12	0.04	0.07	0.10	0.16	0.17	0.29	0.20	0.46	0.16	0.52	0.37	0.93
Good 47	1997.12	0.08	0.10	0.18	0.23	0.33	0.43	0.50	0.80	0.46	1.06	0.76	1.57
Good 48	1997.12	0.03	0.07	0.06	0.17	0.15	0.38	0.40	0.80	0.53	0.95	0.68	1.16
PPP	1987.01	0.10	0.17	0.18	0.33	0.29	0.54	0.38	0.79	0.42	0.96	0.48	1.18

Table 4: Out-of-sample Forecast Evaluation: Clark and West Statistics (US – Japan)

	H_0: **Random Walk with No Drift**						H_0: **Random Walk with Drift**					
Forecast Horizon	3	6	12	24	36	48	3	6	12	24	36	48
Good 01	0.97	0.96	0.68	0.97	2.06	1.48	0.02	0.08	-0.81	-0.17	3.46	2.75
Good 02	0.44	0.76	0.09	1.88	2.98	1.77	-0.32	0.37	-0.10	2.58	3.33	2.94
Good 03	0.68	0.80	0.66	2.78	3.24	2.49	-0.75	-0.33	-0.54	2.63	2.94	2.67
Good 04	0.14	0.14	-0.19	1.10	0.98	2.46	0.50	0.51	-0.43	2.50	3.19	1.86
Good 05	0.19	0.48	0.63	1.99	3.67	3.28	1.83	2.69	2.03	2.41	1.72	1.36
Good 06	1.79	1.98	1.88	2.94	4.36	2.87	1.87	2.15	2.02	2.16	1.91	1.99
Good 07	1.15	1.09	0.82	1.34	1.55	1.79	1.15	1.10	0.69	1.62	3.21	2.81
Good 08	0.08	-0.52	-1.30	-1.39	-0.99	-1.28	-1.19	-0.84	-1.18	-0.84	1.20	2.29
Good 09	0.70	1.01	0.65	2.04	2.13	1.26	-0.71	0.45	0.38	2.30	2.57	2.26
Good 10	0.26	0.27	-0.05	0.59	1.07	1.07	-1.64	-0.52	-0.06	1.10	1.89	1.96
Good 11	0.36	-0.05	-0.90	-0.14	0.04	-0.04	-1.01	-0.03	-0.05	1.29	1.71	2.18
Good 12	0.88	1.00	0.09	1.19	1.08	1.13	0.37	0.99	0.79	1.66	1.69	2.20
Good 13	1.06	1.07	0.81	1.12	1.07	2.14	0.93	0.94	0.60	1.12	1.29	1.96
Good 14	0.89	1.02	0.95	1.46	2.15	2.04	-0.30	0.12	-0.15	1.61	2.48	2.27
Good 15	0.85	0.89	0.47	1.58	2.71	1.41	-0.56	-0.52	-0.94	1.59	2.29	1.95
Good 16	1.26	1.50	1.30	1.38	1.69	1.90	1.36	1.80	1.42	1.89	4.15	2.00
Good 17	0.29	0.03	-0.86	-0.13	0.60	0.26	-1.20	-0.55	-0.85	0.82	1.85	2.32
Good 18	-0.23	-0.49	-0.95	-0.36	-0.14	-0.94	-0.81	-0.28	-0.46	1.08	1.65	1.93
Good 19	0.60	0.56	0.39	1.16	1.83	1.88	-1.91	-1.34	-1.70	0.87	2.98	2.48
Good 20	1.28	1.34	1.20	1.41	1.49	1.84	1.25	1.32	1.16	1.65	3.87	2.26
Good 21	0.71	0.83	0.56	1.32	1.98	2.43	-1.46	-0.58	-0.67	1.12	1.88	1.93
Good 22	0.71	0.85	0.88	1.44	1.81	1.93	-1.16	-0.68	-1.18	1.24	1.88	1.59
Good 23	0.60	0.65	0.72	0.95	1.04	1.99	0.29	0.33	0.36	0.81	1.00	2.47
Good 24	1.32	1.35	1.18	1.32	1.43	1.85	1.29	1.31	1.11	1.63	3.57	2.45
Good 25	1.39	1.77	2.31	4.39	4.36	2.22	0.78	1.30	1.63	2.11	1.77	1.43
Good 26	0.57	0.73	0.56	1.97	2.15	1.49	-1.16	-0.04	0.21	1.64	2.14	1.95
Good 27	0.96	1.07	0.93	1.04	1.12	2.05	0.77	0.88	0.63	0.96	1.54	2.50
Good 28	-0.72	-0.60	-0.61	0.45	1.41	1.49	-0.33	0.08	0.33	1.58	2.59	2.49
Good 29	1.27	1.23	1.09	1.24	1.34	1.95	1.36	1.32	1.06	1.53	3.66	2.47
Good 30	0.95	1.11	0.65	2.42	2.57	1.33	0.80	1.34	1.38	2.18	2.10	2.26
Good 31	1.35	1.36	1.17	1.36	1.46	1.88	1.54	1.54	1.21	1.86	4.60	2.41
Good 32	1.67	1.67	1.43	1.56	1.89	1.64	1.87	1.85	1.54	2.22	5.07	2.12
Good 33	0.51	0.80	0.33	1.43	1.35	0.71	-0.08	0.65	0.81	2.28	2.51	2.22
Good 34	1.25	1.13	0.90	0.94	1.08	2.02	1.28	1.18	0.71	1.07	2.33	2.48
Good 35	0.35	0.33	0.18	0.64	0.73	2.22	-0.11	-0.14	-0.48	0.30	0.05	-0.59
Good 36	1.65	2.04	2.27	3.19	2.83	2.16	1.37	1.89	2.36	3.24	2.84	2.43
Good 37	0.80	0.98	0.60	1.74	2.87	1.53	-0.28	0.28	-0.17	1.94	2.61	2.14
Good 38	1.06	1.56	1.45	1.74	1.93	1.61	0.67	1.30	1.33	2.05	2.28	2.08
Good 39	0.13	0.27	-0.19	1.30	3.27	3.29	-1.72	-1.14	-0.68	1.33	2.04	2.11
Good 40	1.32	1.37	1.20	1.34	1.41	1.83	1.39	1.48	1.22	1.71	4.05	2.47
Good 41	0.58	0.69	0.61	0.97	0.90	2.15	0.41	0.56	0.25	1.09	0.90	1.98
Good 42	0.43	0.53	0.66	1.43	1.69	2.08	0.71	0.84	0.73	2.25	4.33	2.31
Good 43	-0.04	0.13	-0.29	0.44	0.08	-0.05	0.84	1.28	1.30	1.97	2.27	2.30
Good 44	0.30	0.47	0.64	1.34	1.62	1.75	-0.03	0.10	0.25	1.34	2.64	2.35
Good 45	0.64	0.67	0.77	1.03	1.01	1.99	0.38	0.40	0.45	0.95	1.05	2.77
Good 46	-1.22	-0.72	-0.83	-0.14	-0.37	-2.26	-0.38	0.12	-0.18	0.59	0.89	1.87
Good 47	0.03	0.02	-0.22	0.23	0.39	2.15	-0.69	-0.59	-1.20	-0.30	-0.57	-1.30
Good 48	-0.80	-0.76	-1.41	-0.32	0.59	0.69	0.27	0.85	0.88	2.24	3.04	3.02
Good 49	-0.95	-0.56	-0.74	-0.03	-0.20	-2.49	-0.08	0.32	-0.01	0.64	0.95	1.86
Good 50	0.30	0.88	0.27	2.07	0.38	-2.16	1.13	1.55	0.91	1.81	1.30	1.91
Good 51	0.75	1.11	1.04	1.82	2.06	1.64	1.00	1.40	1.10	3.05	4.88	2.12
Good 52	-0.98	-1.05	-1.35	-1.43	-0.97	-0.84	-1.11	-1.03	-1.51	-1.55	-0.68	0.49
Good 53	-1.28	-1.08	-1.20	-1.05	-0.57	-1.18	-0.88	-0.44	-0.60	1.14	1.86	1.88
Good 54	-0.68	-0.37	-0.44	0.75	0.92	0.67	0.22	0.61	0.85	2.10	2.44	2.37
Good 55	-1.07	-1.09	-1.23	-1.29	-0.86	-1.68	-1.08	-0.91	-1.08	-0.12	1.50	1.45
Good 56	-1.31	-1.20	-1.30	-1.10	-0.73	-0.93	-0.81	-0.29	-0.42	1.37	2.09	2.29
Good 57	-0.81	-1.20	-1.40	-1.18	-0.82	-0.47	-0.67	-0.06	-0.18	1.39	2.23	2.61
Good 58	1.28	1.24	1.02	1.18	1.27	2.00	1.31	1.25	0.90	1.43	3.03	2.47
Good 59	-0.15	0.18	0.03	1.85	1.49	0.89	-0.04	0.65	0.92	2.67	2.76	2.25
Good 60	0.17	0.16	-0.37	1.69	4.30	1.78	-2.23	-1.56	-1.80	0.62	1.54	1.67
Good 61	0.16	0.13	-0.64	0.74	1.35	0.68	-1.41	-0.66	-0.86	1.38	2.63	2.19
Good 62	0.19	0.07	-1.09	-0.35	0.77	-0.48	-1.22	-0.89	-1.20	0.66	2.24	1.84
Good 63	0.62	0.84	1.09	1.83	1.84	1.55	-0.79	-0.30	-0.04	0.50	0.12	-0.41
Good 64	-0.15	0.02	-0.13	0.76	1.25	2.16	-0.64	-0.57	-0.88	0.33	1.24	2.14
Good 65	0.48	0.51	0.48	0.94	1.05	2.03	0.11	0.13	0.00	0.76	0.85	1.87
Good 66	1.21	1.17	0.17	1.84	1.79	1.06	1.06	1.15	1.01	2.19	2.79	2.33
Good 67	1.18	1.47	1.73	2.65	2.33	1.65	0.28	0.67	0.72	1.45	1.13	0.89
PPP	0.62	0.87	0.36	1.78	1.56	0.89	-0.30	0.58	0.65	2.44	2.60	2.23

Table 5: Out-of-sample Forecast Evaluation: Clark and West Statistics (US – UK)

	H_0: **Random Walk with No Drift**						H_0: **Random Walk with Drift**					
Forecast Horizon	3	6	12	24	36	48	3	6	12	24	36	48
Good 01	0.46	0.73	0.74	1.99	1.64	1.88	0.49	0.83	0.89	1.94	1.42	1.68
Good 02	0.37	0.63	0.54	0.73	0.65	1.03	0.62	0.87	0.95	1.33	1.32	1.49
Good 03	1.47	1.39	1.59	2.16	1.53	2.50	1.54	1.44	1.66	2.37	1.66	2.68
Good 04	0.66	0.73	0.30	0.70	0.90	1.71	0.91	0.96	0.70	0.93	0.96	1.85
Good 05	1.02	0.01	-0.57	1.04	0.04	0.35	1.25	0.92	0.92	1.33	0.85	1.42
Good 06	-1.34	-1.17	-1.60	0.24	-2.34	4.37	-1.21	0.16	-2.46	2.37	-0.86	2.16
Good 07	-0.28	0.14	0.20	1.38	1.62	2.10	-0.04	0.35	0.49	1.65	1.81	2.26
Good 08	1.16	1.38	1.38	-0.56	0.07	-2.41	1.17	1.30	1.30	0.20	0.59	0.07
Good 09	0.46	0.72	0.87	1.99	1.79	2.49	0.73	0.91	1.13	2.17	2.11	2.70
Good 10	1.75	1.97	1.87	1.13	-0.14	-0.80	1.83	2.03	2.05	1.56	1.00	0.40
Good 11	0.09	0.11	-0.08	0.07	0.53	1.59	0.28	0.33	0.19	0.31	0.58	1.59
Good 12	0.53	1.31	1.39	1.17	-1.80	1.02	1.21	1.49	1.59	1.63	0.15	1.54
Good 13	0.93	1.00	1.18	2.08	1.78	2.55	1.05	1.10	1.33	2.56	1.95	2.74
Good 14	-0.68	-0.69	-0.93	-0.70	-0.11	1.21	-0.59	-0.46	-0.67	-0.27	0.20	1.48
Good 15	0.00	0.81	0.68	1.60	0.54	1.13	0.07	0.98	0.93	1.91	0.70	2.33
Good 16	-0.57	-0.35	-0.55	-0.23	0.32	1.37	-0.27	0.26	-0.06	0.17	0.45	1.54
Good 17	-0.22	0.01	0.12	1.48	1.80	2.65	-0.04	0.19	0.37	1.76	1.82	2.68
Good 18	-0.48	-0.01	0.66	1.91	0.69	1.52	-0.07	0.52	1.15	2.22	1.18	1.92
Good 19	1.16	1.22	1.56	2.37	-0.76	-0.69	1.59	1.36	1.58	1.87	-1.05	-0.26
Good 20	0.80	0.78	0.57	0.17	-0.42	-1.30	1.13	1.08	1.03	0.82	0.30	-0.19
Good 21	-1.22	-2.33	-2.29	-0.64	-0.62	1.71	-0.53	-0.26	-1.86	1.16	0.50	1.81
Good 22	0.95	1.10	1.62	2.59	2.23	3.00	1.14	1.28	1.76	2.80	2.43	3.11
Good 23	1.12	1.12	1.23	1.32	0.72	-0.54	1.39	1.36	1.50	1.75	1.30	0.51
Good 24	-0.63	-0.14	0.95	2.48	1.78	2.55	-0.44	0.13	1.19	2.50	1.68	2.40
Good 25	0.05	1.17	1.22	0.80	0.74	1.64	0.46	1.23	1.31	1.27	1.08	1.84
Good 26	-0.28	-0.21	-0.33	-0.19	-0.63	-0.46	0.12	0.19	0.19	0.48	0.13	0.27
Good 27	-1.32	-0.97	-0.75	-0.64	-4.58	0.62	-1.31	-0.86	-0.35	0.77	-2.57	1.42
Good 28	0.08	0.63	0.52	-0.24	-0.57	0.82	0.88	1.08	1.29	0.93	0.52	1.51
Good 29	-0.82	-0.92	-1.81	0.02	-1.40	1.93	-1.06	-1.21	-2.46	1.74	-0.29	1.98
Good 30	0.12	0.80	0.59	0.40	-0.71	0.79	0.98	1.15	1.32	1.21	0.47	1.50
Good 31	-1.62	-1.70	-2.34	-1.09	-0.54	1.10	-1.56	-1.78	-3.27	1.09	0.52	1.62
Good 32	-0.24	0.10	0.07	1.37	1.54	2.06	-0.22	0.24	0.28	1.55	1.43	1.98
Good 33	0.52	0.68	0.74	1.51	1.35	2.15	0.70	0.83	0.97	1.81	1.43	2.26
Good 34	0.91	1.09	1.39	2.07	1.77	2.33	1.06	1.22	1.59	2.23	1.73	2.40
Good 35	0.54	0.77	1.00	1.44	0.93	2.19	0.80	0.97	1.25	1.92	1.30	2.40
Good 36	-0.18	0.46	0.61	2.00	1.54	2.23	-0.09	0.65	0.92	2.17	1.51	2.33
Good 37	-0.65	-0.55	-0.03	0.30	-0.29	-1.41	-0.45	-0.24	0.44	0.61	0.09	-0.46
Good 38	-0.32	-0.25	-0.30	-0.14	-0.47	-0.84	0.11	0.19	0.24	0.49	0.25	0.00
Good 39	0.82	0.86	0.72	-0.23	-1.63	-2.68	1.09	1.09	0.99	0.24	-0.98	-1.56
Good 40	0.37	0.44	0.23	-0.11	-0.87	-1.21	0.73	0.78	0.65	0.44	-0.13	-0.32
Good 41	-0.40	-0.41	-0.64	-0.66	-1.11	-1.83	0.03	0.04	-0.17	-0.12	-0.56	-0.90
Good 42	-0.24	0.30	0.39	1.55	1.41	2.32	-0.21	0.43	0.58	1.67	1.36	2.41
Good 43	-0.97	-0.93	-1.20	-0.90	0.27	1.38	-1.18	-1.09	-1.39	-1.08	0.13	1.31
Good 44	0.79	1.44	1.90	2.43	-0.73	-0.47	1.30	1.49	1.67	1.86	-1.13	-0.08
Good 45	-0.82	-0.78	-1.04	-1.16	-1.55	-2.50	-0.65	-0.63	-0.88	-0.88	-1.27	-1.93
Good 46	-1.61	-1.71	-2.25	-0.49	-0.33	0.84	-1.23	-1.54	-2.59	1.03	0.62	1.50
Good 47	-0.04	0.18	0.58	0.80	0.00	1.47	0.77	0.89	1.19	1.23	0.92	1.72
Good 48	-1.68	-0.56	-0.56	0.19	-0.32	1.17	-0.62	0.89	0.96	1.08	0.67	1.73
PPP	1.02	1.06	1.43	1.94	1.29	1.81	1.32	1.31	1.67	2.37	1.87	2.06

Table 6: The Ratio of RMSPE for RW with No Drift to RMSPE for RW with Drift (US – Japan)

Starting Date	Date of First Forecast	3-month	6-month	12-month	24-month	36-month	48-month
1973M03	1983M01	1.003	1.004	1.006	0.997	0.987	0.983
1977M12	1983M01	1.000	0.998	0.993	0.970	0.944	0.923
1978M01	1983M01	1.000	0.998	0.993	0.971	0.945	0.925
1980M01	1994M10	0.980	0.955	0.893	0.830	0.724	0.509
1997M12	2003M10	0.994	0.983	0.989	0.911	0.857	1.010

Table 7: SPA Test Groups (US – Japan)

Case	Group of Goods	Observations	Sample used for out-of-sample analysis	Date of First Forecast
1	25 goods	first observation at 1973M03	All available	1983M01
2	13 goods	first observation at 1977M12	All available	1983M01
3	26 goods	first observation at 1997M12	All available	2003M10
4	38 goods	group 1 and 2	All available	1983M01
5	64 goods	group 1, 2 and 3	All available	2003M10
6	64 goods	group 1, 2 and 3	from 1997M12 and onwards	2003M10

Table 8: Tests for Superior Predictive Ability: Individual Groups (US – Japan)

			H_0: **RW with No Drift**					H_0: **RW with Drift**				
	k		Stat.	p-value	10% CV	5% CV	1% CV	Stat.	p-value	10% CV	5% CV	1% CV
Case 1:	3	Lower	1.897	0.130	2.042	2.375	2.948	1.523	0.323	2.096	2.388	2.954
		Consistent		0.169	2.145	2.475	3.093		0.502	2.359	2.612	3.107
		Upper		0.169	2.145	2.475	3.093		0.513	2.366	2.618	3.107
	6	Lower	2.734	0.026	2.131	2.467	3.097	2.411	0.064	2.226	2.505	3.102
		Consistent		0.030	2.204	2.521	3.147		0.086	2.354	2.621	3.201
		Upper		0.030	2.204	2.521	3.147		0.087	2.356	2.624	3.201
	12	Lower	3.167	0.008	2.083	2.406	3.053	3.024	0.009	2.047	2.378	2.981
		Consistent		0.009	2.162	2.486	3.136		0.014	2.273	2.570	3.133
		Upper		0.009	2.169	2.498	3.139		0.014	2.301	2.585	3.141
	24	Lower	5.194	0.000	2.205	2.521	3.193	5.090	0.000	2.217	2.547	3.247
		Consistent		0.000	2.209	2.526	3.193		0.000	2.256	2.575	3.249
		Upper		0.000	2.209	2.526	3.193		0.000	2.256	2.575	3.249
	36	Lower	7.041	0.000	2.175	2.514	3.187	6.090	0.000	2.190	2.544	3.245
		Consistent		0.000	2.175	2.514	3.187		0.000	2.190	2.544	3.245
		Upper		0.000	2.175	2.514	3.187		0.000	2.190	2.544	3.245
	48	Lower	6.301	0.000	2.136	2.522	3.271	6.205	0.000	2.084	2.481	3.322
		Consistent		0.000	2.144	2.530	3.274		0.000	2.141	2.524	3.348
		Upper		0.000	2.144	2.530	3.274		0.000	2.141	2.524	3.348
Case 2:	3	Lower	2.104	0.066	1.881	2.233	2.903	1.930	0.085	1.848	2.177	2.783
		Consistent		0.088	2.047	2.351	2.960		0.147	2.089	2.366	2.892
		Upper		0.088	2.047	2.351	2.960		0.147	2.089	2.366	2.892
	6	Lower	2.642	0.018	1.867	2.212	2.865	2.528	0.023	1.847	2.174	2.852
		Consistent		0.024	2.046	2.345	2.963		0.024	1.957	2.249	2.865
		Upper		0.024	2.046	2.345	2.963		0.024	1.957	2.249	2.865
	12	Lower	2.907	0.004	1.612	1.994	2.634	2.889	0.008	1.745	2.078	2.784
		Consistent		0.006	1.742	2.113	2.745		0.009	1.825	2.144	2.823
		Upper		0.008	1.992	2.294	2.846		0.009	1.917	2.216	2.823
	24	Lower	3.742	0.001	1.820	2.161	2.816	3.603	0.003	1.850	2.239	3.022
		Consistent		0.001	1.994	2.312	2.991		0.003	1.936	2.306	3.045
		Upper		0.001	2.082	2.401	3.001		0.003	2.008	2.337	3.045
	36	Lower	6.217	0.000	1.940	2.268	2.854	5.177	0.000	1.875	2.266	3.032
		Consistent		0.000	2.045	2.370	2.956		0.000	1.954	2.312	3.048
		Upper		0.000	2.111	2.417	2.980		0.000	1.954	2.312	3.048
	48	Lower	6.821	0.000	1.828	2.183	3.009	6.651	0.000	1.940	2.303	3.015
		Consistent		0.000	1.872	2.215	3.027		0.000	1.940	2.303	3.015
		Upper		0.000	2.069	2.386	3.069		0.000	1.940	2.303	3.015
Case 3:	3	Lower	2.018	0.105	2.039	2.355	3.039	2.167	0.084	2.086	2.430	3.106
		Consistent		0.123	2.120	2.421	3.051		0.094	2.143	2.464	3.133
		Upper		0.123	2.120	2.421	3.051		0.094	2.143	2.464	3.133
	6	Lower	2.310	0.050	1.992	2.319	2.948	2.472	0.036	2.034	2.339	2.995
		Consistent		0.056	2.040	2.357	2.968		0.039	2.081	2.380	3.001
		Upper		0.056	2.040	2.357	2.968		0.039	2.081	2.380	3.001
	12	Lower	2.565	0.026	1.940	2.259	2.982	2.503	0.033	1.990	2.326	3.048
		Consistent		0.033	2.068	2.389	3.048		0.038	2.072	2.398	3.081
		Upper		0.033	2.070	2.390	3.048		0.040	2.093	2.407	3.085
	24	Lower	3.363	0.004	1.930	2.240	2.966	4.539	0.000	1.967	2.289	3.024
		Consistent		0.004	1.940	2.252	2.967		0.000	1.982	2.299	3.024
		Upper		0.004	1.940	2.252	2.967		0.000	1.982	2.299	3.024
	36	Lower	3.224	0.004	1.815	2.164	2.800	6.047	0.000	2.043	2.384	3.040
		Consistent		0.004	1.852	2.177	2.812		0.000	2.057	2.393	3.040
		Upper		0.004	1.852	2.177	2.812		0.000	2.057	2.393	3.040
	48	Lower	3.998	0.001	1.646	2.055	2.823	3.316	0.005	1.854	2.253	3.037
		Consistent		0.001	1.646	2.055	2.823		0.006	1.981	2.360	3.118
		Upper		0.001	1.947	2.290	2.977		0.007	2.037	2.411	3.138

Table 9: Tests for Superior Predictive Ability: Mixing Groups (US – Japan)

			H_0: **RW with No Drift**					H_0: **RW with Drift**				
	k		Stat.	p-value	10% CV	5% CV	1% CV	Stat.	p-value	10% CV	5% CV	1% CV
Case 4:	3	Lower	2.104	0.103	2.119	2.445	3.074	1.712	0.257	2.173	2.442	3.040
		Consistent		0.137	2.258	2.552	3.176		0.425	2.420	2.663	3.183
		Upper		0.137	2.258	2.552	3.176		0.437	2.429	2.667	3.183
	6	Lower	2.734	0.029	2.195	2.522	3.152	2.411	0.071	2.262	2.551	3.123
		Consistent		0.036	2.285	2.595	3.194		0.095	2.388	2.658	3.238
		Upper		0.036	2.285	2.595	3.194		0.096	2.393	2.660	3.238
	12	Lower	3.167	0.009	2.119	2.454	3.128	3.024	0.010	2.097	2.436	3.028
		Consistent		0.010	2.197	2.536	3.169		0.015	2.304	2.602	3.175
		Upper		0.010	2.265	2.576	3.172		0.015	2.338	2.622	3.180
	24	Lower	5.194	0.000	2.270	2.596	3.282	5.090	0.000	2.288	2.627	3.362
		Consistent		0.000	2.331	2.676	3.314		0.000	2.326	2.651	3.372
		Upper		0.000	2.371	2.688	3.340		0.000	2.363	2.664	3.372
	36	Lower	7.041	0.000	2.325	2.656	3.273	6.090	0.000	2.300	2.646	3.398
		Consistent		0.000	2.378	2.697	3.297		0.000	2.300	2.646	3.398
		Upper		0.000	2.398	2.714	3.313		0.000	2.321	2.651	3.398
	48	Lower	6.821	0.000	2.219	2.597	3.386	6.205	0.000	2.219	2.580	3.393
		Consistent		0.000	2.231	2.604	3.386		0.000	2.283	2.615	3.414
		Upper		0.000	2.329	2.661	3.402		0.000	2.283	2.615	3.414
Case 5:	3	Lower	1.712	0.257	2.173	2.442	3.040	2.463	0.071	2.286	2.632	3.355
		Consistent		0.425	2.420	2.663	3.183		0.084	2.374	2.697	3.402
		Upper		0.437	2.429	2.667	3.183		0.084	2.374	2.697	3.402
	6	Lower	2.310	0.067	2.124	2.439	3.049	3.416	0.005	2.286	2.623	3.245
		Consistent		0.079	2.209	2.497	3.085		0.006	2.319	2.646	3.253
		Upper		0.079	2.209	2.497	3.085		0.006	2.319	2.646	3.253
	12	Lower	3.068	0.013	2.087	2.430	3.225	4.131	0.001	2.230	2.587	3.224
		Consistent		0.014	2.188	2.515	3.228		0.001	2.285	2.621	3.256
		Upper		0.014	2.195	2.517	3.228		0.001	2.300	2.624	3.259
	24	Lower	4.886	0.000	2.126	2.462	3.171	5.051	0.000	2.209	2.537	3.222
		Consistent		0.000	2.138	2.464	3.171		0.000	2.217	2.542	3.230
		Upper		0.000	2.138	2.464	3.171		0.000	2.217	2.542	3.230
	36	Lower	4.411	0.000	2.109	2.424	3.068	6.131	0.000	2.220	2.545	3.149
		Consistent		0.000	2.124	2.435	3.068		0.000	2.265	2.579	3.187
		Upper		0.000	2.124	2.435	3.068		0.000	2.265	2.579	3.187
	48	Lower	4.145	0.001	1.978	2.311	2.989	3.394	0.006	2.020	2.390	3.132
		Consistent		0.001	2.066	2.364	3.037		0.007	2.125	2.486	3.195
		Upper		0.001	2.186	2.486	3.127		0.007	2.239	2.544	3.232
Case 6:	3	Lower	3.026	0.022	2.290	2.646	3.311	3.292	0.014	2.335	2.673	3.396
		Consistent		0.025	2.396	2.720	3.376		0.014	2.405	2.712	3.406
		Upper		0.025	2.396	2.720	3.376		0.015	2.449	2.744	3.426
	6	Lower	3.354	0.009	2.302	2.624	3.295	3.889	0.002	2.403	2.725	3.321
		Consistent		0.009	2.361	2.653	3.305		0.002	2.427	2.738	3.328
		Upper		0.009	2.361	2.653	3.305		0.002	2.427	2.738	3.328
	12	Lower	4.312	0.001	2.243	2.581	3.330	3.884	0.002	2.316	2.670	3.311
		Consistent		0.001	2.309	2.623	3.330		0.003	2.376	2.706	3.354
		Upper		0.001	2.352	2.658	3.332		0.003	2.412	2.725	3.379
	24	Lower	6.044	0.000	2.240	2.573	3.264	6.133	0.000	2.350	2.662	3.367
		Consistent		0.000	2.244	2.576	3.264		0.000	2.362	2.680	3.367
		Upper		0.000	2.244	2.576	3.264		0.000	2.362	2.680	3.367
	36	Lower	4.981	0.000	2.179	2.506	3.153	6.047	0.000	2.337	2.659	3.328
		Consistent		0.000	2.194	2.518	3.153		0.000	2.341	2.665	3.333
		Upper		0.000	2.194	2.518	3.153		0.000	2.341	2.665	3.333
	48	Lower	4.223	0.000	1.921	2.281	2.989	3.888	0.002	2.087	2.427	3.182
		Consistent		0.000	2.119	2.431	3.082		0.002	2.161	2.493	3.221
		Upper		0.001	2.244	2.538	3.208		0.002	2.200	2.539	3.248

Table 10: A List of Goods that Display Statistical Significance in Exchange Rate Predictability (US – Japan)

Forecast Horizon	Good No.	Good Description	Forecast Horizon	Good No.	Good Description
3-month	Good 06	Frozen fish and seafood	24-month	Good 02	Pork chops
	Good 07	Fresh fish and seafood		Good 05	Ham
	Good 20	Juices and non-alcoholic drinks		Good 06	Frozen fish and seafood
	Good 24	Canned vegetables		Good 30	Spices, seasonings, condiments, sauces
	Good 29	Other beverage materials including tea		Good 33	Rent of primary residence
	Good 31	Full service meals and snacks		Good 36	Electricity
	Good 32	Food at employee sites and schools		Good 38	Utility (piped) gas service
	Good 36	Electricity		Good 59	New vehicles
	Good 40	Domestic services		Good 66	Intracity transportation
	Good 58	Laundry and dry cleaning services			
6-month	Good 06	Frozen fish and seafood	36-month	Good 02	Pork chops
	Good 30	Spices, seasonings, condiments, sauces		Good 06	Frozen fish and seafood
	Good 32	Food at employee sites and schools		Good 09	Fresh whole milk
	Good 36	Electricity		Good 15	White bread
	Good 38	Utility (piped) gas service		Good 30	Spices, seasonings, condiments, sauces
	Good 66	Intracity transportation		Good 36	Electricity
				Good 38	Utility (piped) gas service
				Good 66	Intracity transportation
12-month	Good 06	Frozen fish and seafood	48-month	Good 14	Eggs
	Good 36	Electricity		Good 19	Bananas
				Good 36	Electricity

Figure 1: In-sample Regression *p*-value (US – Japan)

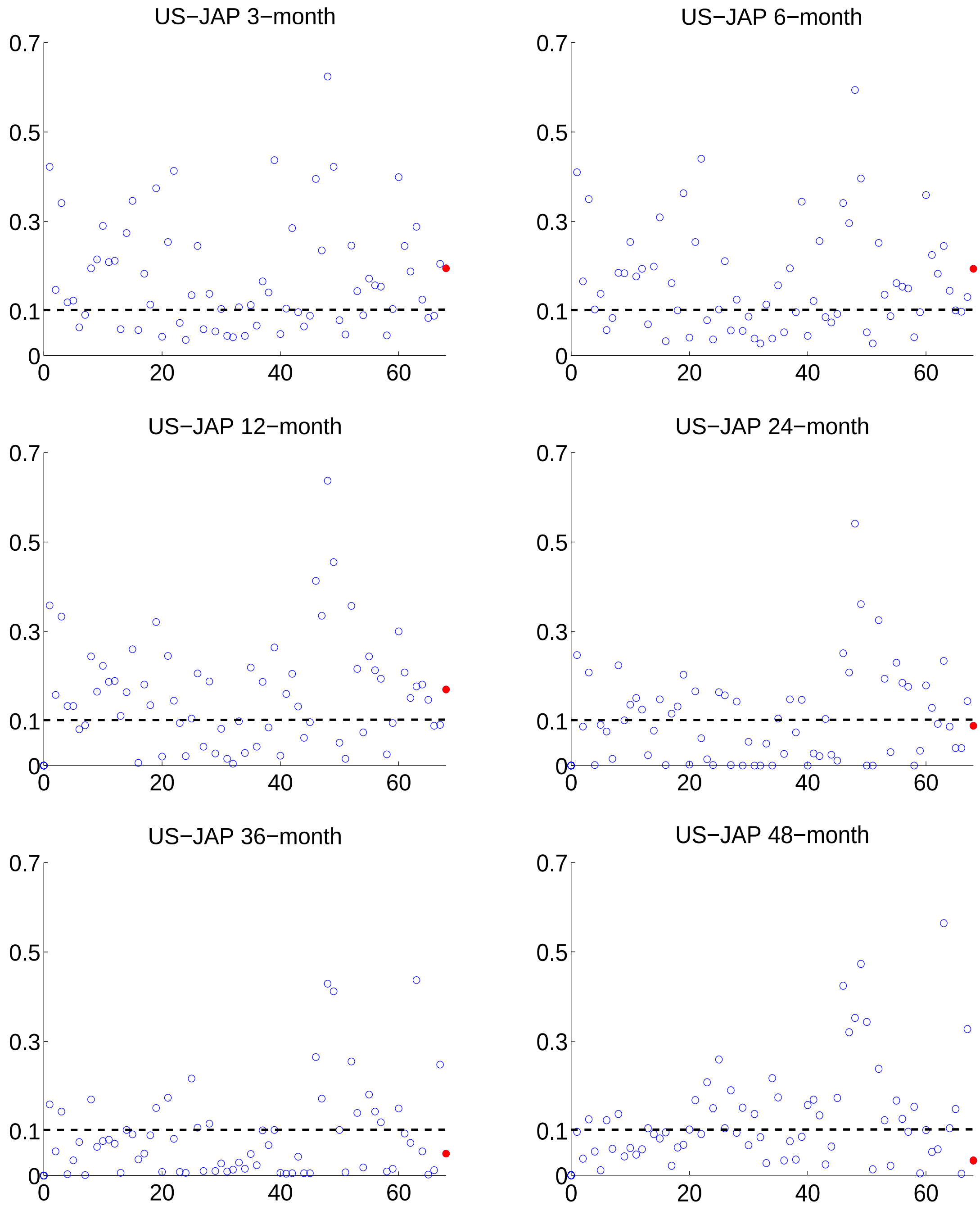


Figure 2: In-sample Regression p-value (US – UK)

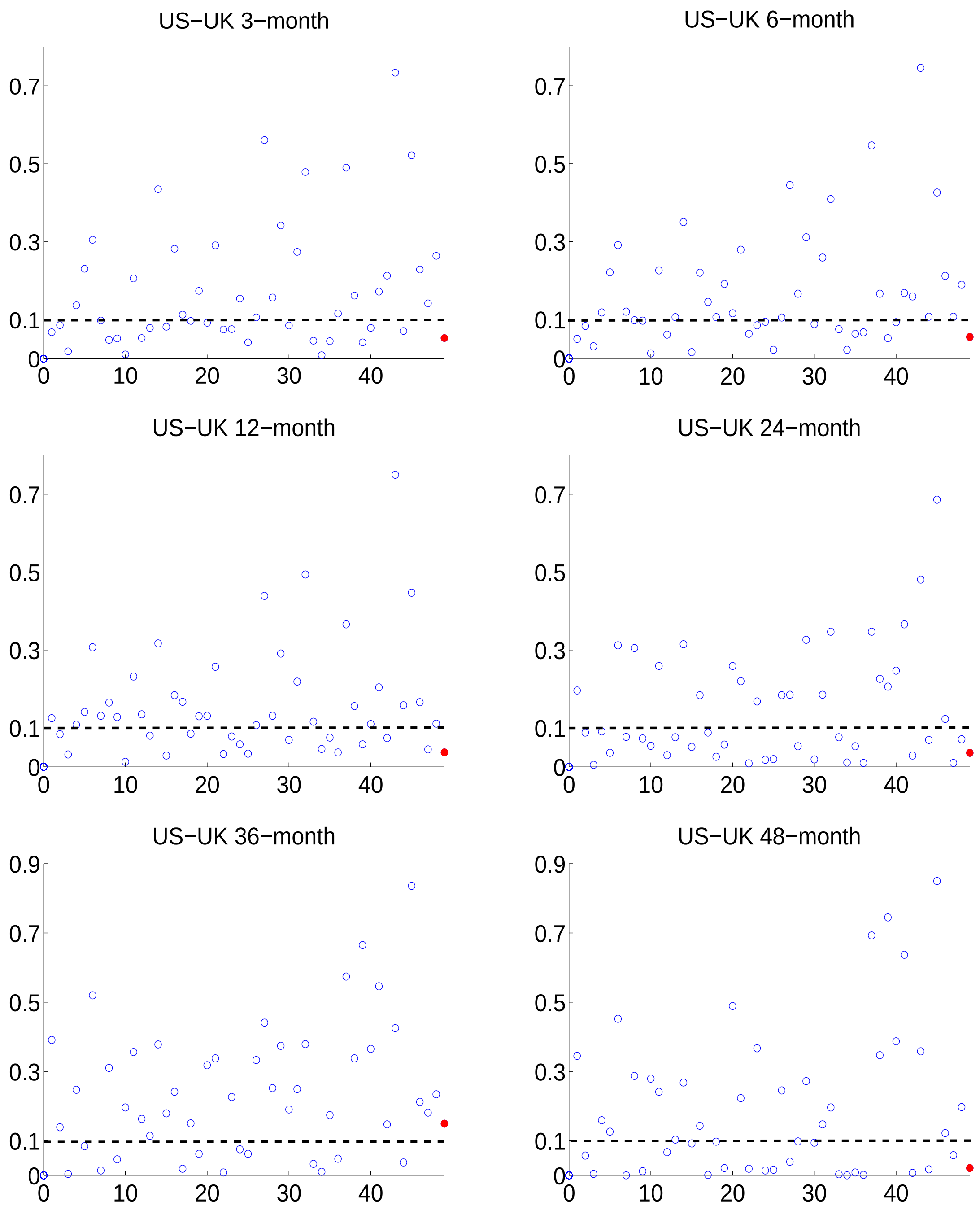
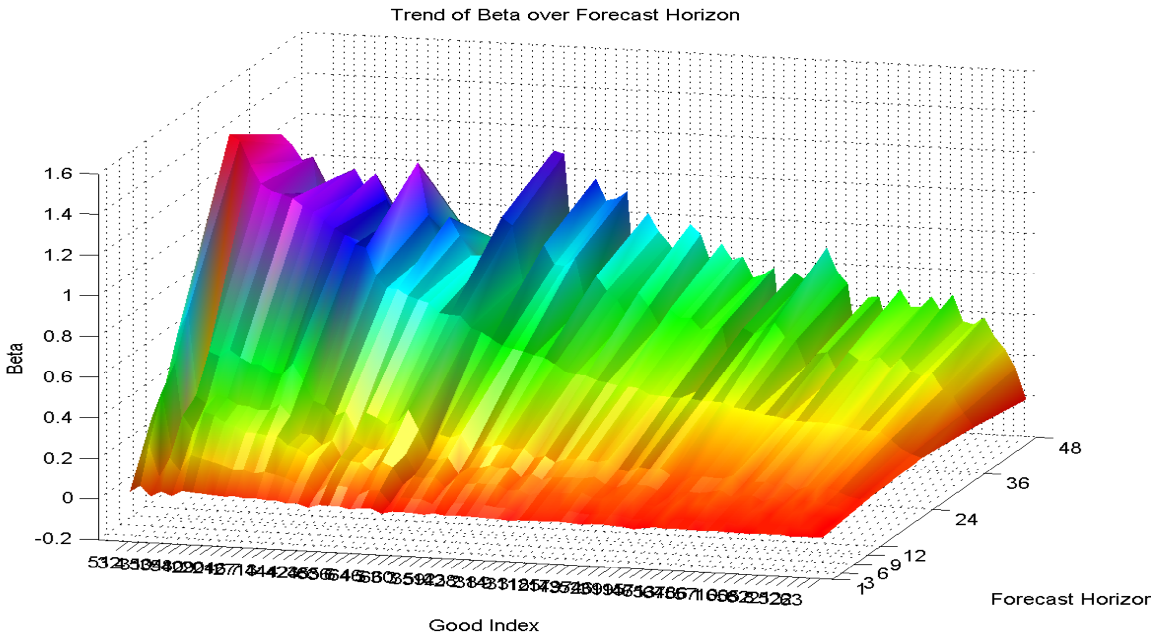


Figure 3: In-sample Regression Coefficient β and R-square over Forecast Horizon (US – Japan)



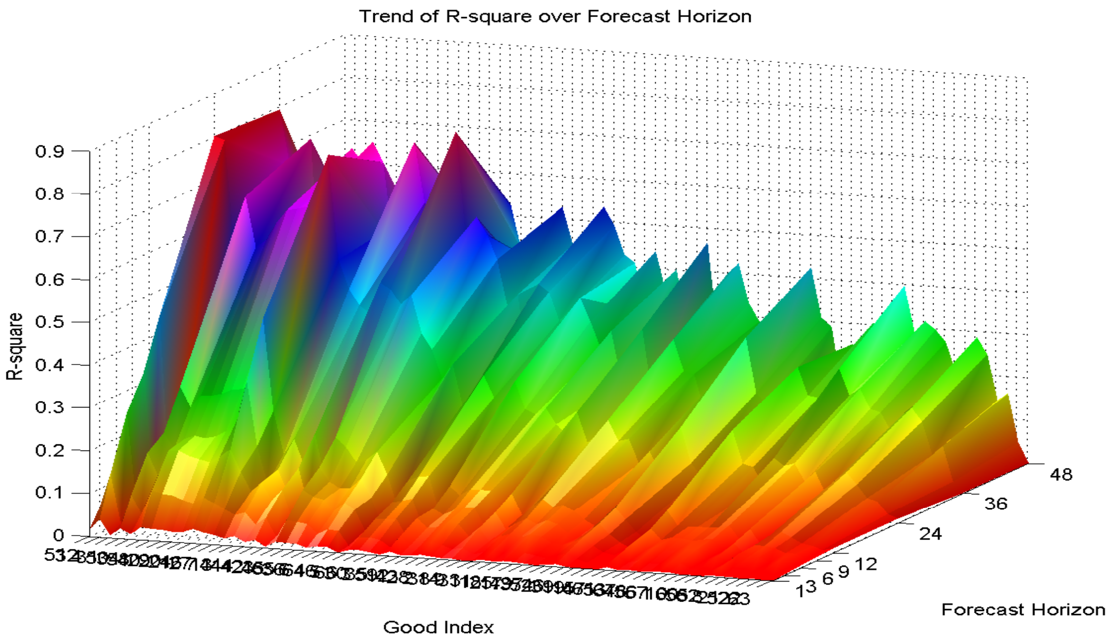


Figure 4: Out-of-sample CW p-value (US – Japan)

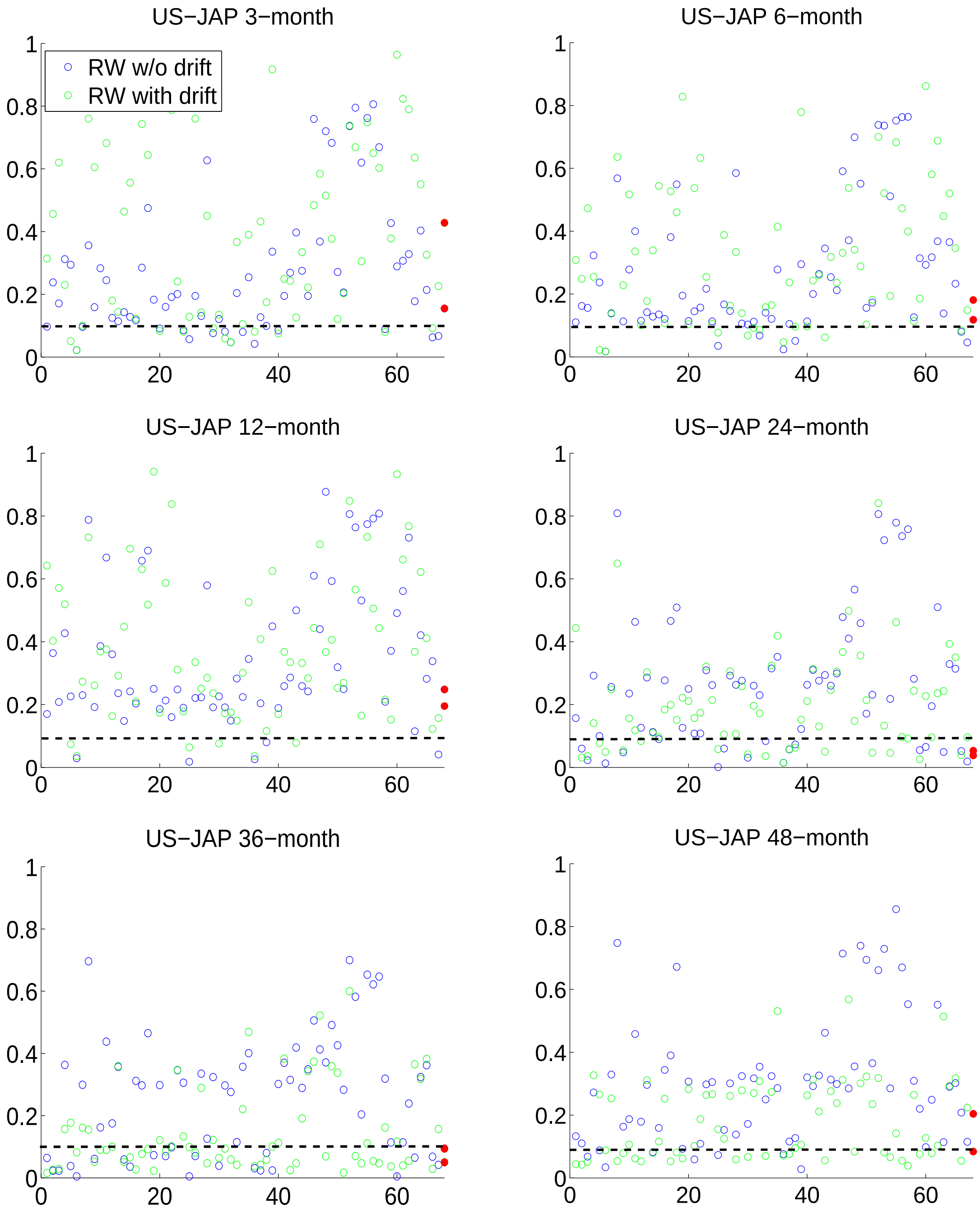
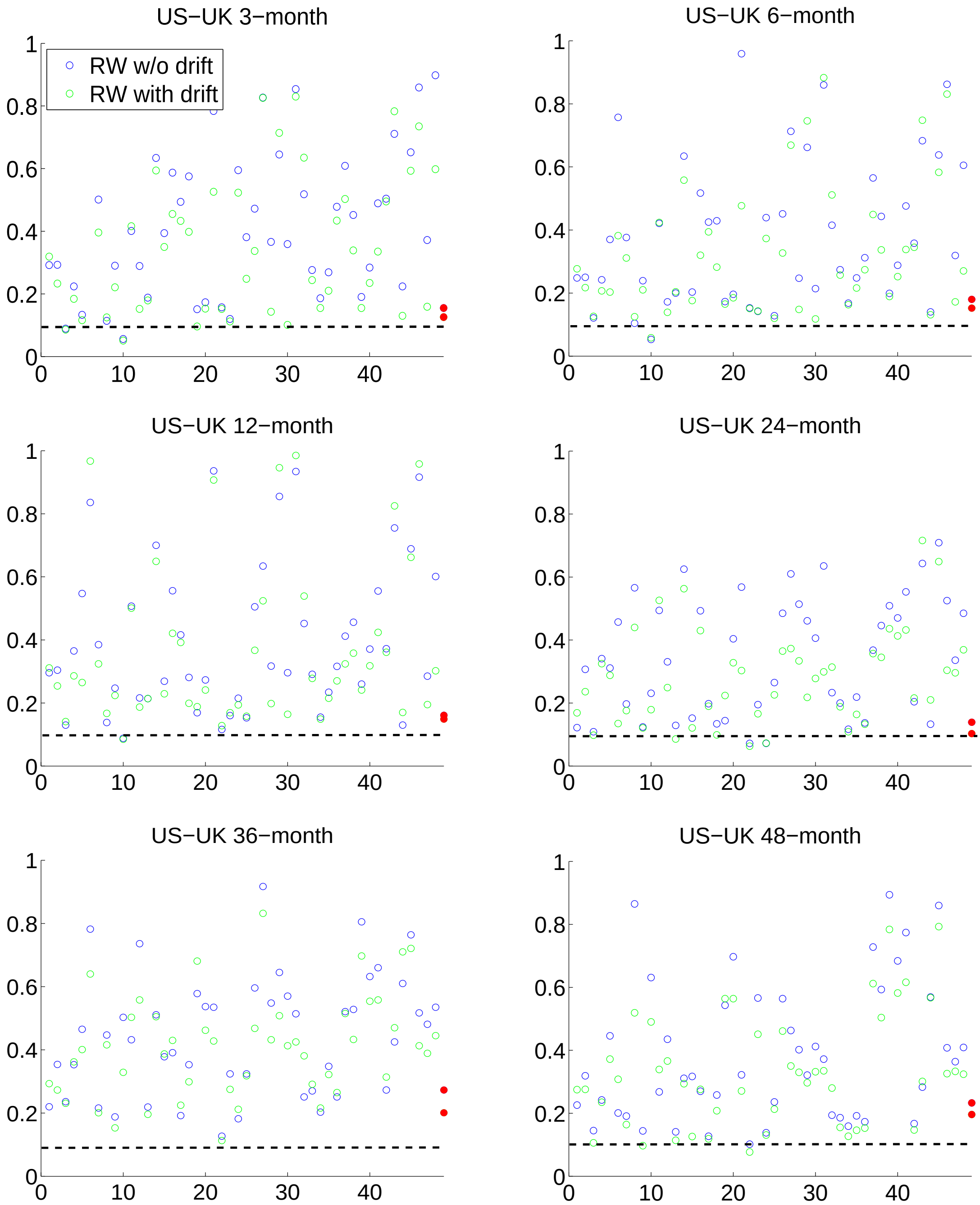


Figure 5: Out-of-sample CW p-value (US – UK)



A Appendix: Description of Data Coverage

Table A.1: Data coverage: US – Japan

US – Japan

Good No.	Good Description	Good No.	Good Description
Good 01	Beef and veal	Good 35	Repair of household goods
Good 02	Pork chops	Good 36	Electricity
Good 03	Poultry	Good 37	Water and sewerage maintenance
Good 04	Bacon, breakfast sausage, and related products	Good 38	Utility (piped) gas service
Good 05	Ham	Good 39	Fuel oil and other fuels
Good 06	Frozen fish and seafood	Good 40	Domestic services
Good 07	Fresh fish and seafood	Good 41	Household cleaning products
Good 08	Canned fish and seafood	Good 42	Household paper products
Good 09	Fresh whole milk	Good 43	Bedroom furniture
Good 10	Butter	Good 44	Floor coverings
Good 11	Cheese and related products	Good 45	Window coverings
Good 12	Ice cream and related products	Good 46	Other linens
Good 13	Other dairy and related products	Good 47	Major appliances
Good 14	Eggs	Good 48	Clocks, lamps, and decorator goods
Good 15	White bread	Good 49	Dishes and flatware
Good 16	Fresh biscuits, rolls, muffins	Good 50	Nonelectric cookware and tableware
Good 17	Rice, pasta, cornmeal	Good 51	Tools, hardware and supplies
Good 18	Flour and prepared flour mixes	Good 52	Women's apparel
Good 19	Bananas	Good 53	Men's apparel
Good 20	Juices and non-alcoholic drinks	Good 54	Infants' and toddlers' apparel
Good 21	Tomatoes	Good 55	Women's footwear
Good 22	Lettuce	Good 56	Men's footwear
Good 23	Canned fruits	Good 57	Boys' and girls' footwear
Good 24	Canned vegetables	Good 58	Laundry and dry cleaning services
Good 25	Sugar and sweets	Good 59	New vehicles
Good 26	Margarine	Good 60	Gasoline (all types)
Good 27	Other fats and oils including peanut butter	Good 61	Tires
Good 28	Coffee	Good 62	Motor vehicle maintenance and repair
Good 29	Other beverage materials including tea	Good 63	Motor vehicle insurance
Good 30	Spices, seasonings, condiments, sauces	Good 64	State and local registration and license
Good 31	Full service meals and snacks	Good 65	Parking and other fees
Good 32	Food at employee sites and schools	Good 66	Intracity transportation
Good 33	Rent of primary residence	Good 67	Airline fare
Good 34	Tenants' and household insurance		

Table A.2: Data coverage: US – UK

US – UK

Good No.	Good Description	Good No.	Good Description
Good 01	Beef and veal	Good 25	Fuel oil and other fuels
Good 02	Pork chops	Good 26	Bedroom furniture
Good 03	Other meats	Good 27	Major appliances
Good 04	Poultry	Good 28	Floor coverings
Good 05	Bacon, breakfast sausage, and related products	Good 29	Other appliances
Good 06	Fresh fish and seafood	Good 30	Miscellaneous household products
Good 07	Fresh whole milk	Good 31	Domestic services
Good 08	Butter	Good 32	Women's apparel
Good 09	Cheese and related products	Good 33	Men's apparel
Good 10	Eggs	Good 34	Infants' and toddlers' apparel
Good 11	White bread	Good 35	Men's footwear
Good 12	Fresh biscuits, rolls, muffins	Good 36	New vehicles
Good 13	Breakfast cereal	Good 37	Gasoline (all types)
Good 14	Fresh Fruits	Good 38	Motor vehicle maintenance and repair
Good 15	Potatoes	Good 39	Motor vehicle insurance
Good 16	Fresh vegetables	Good 40	Intracity transportation
Good 17	Sugar and sweets	Good 41	Personal care services
Good 18	Coffee	Good 42	Sporting goods
Good 19	Other beverage materials including tea	Good 43	Audio equipment
Good 20	Rent of primary residence	Good 44	Audio discs, tapes and other media
Good 21	Repair of household items	Good 45	Televisions
Good 22	Electricity	Good 46	Newspapers and magazines
Good 23	Water and sewerage maintenance	Good 47	Postage and delivery services
Good 24	Utility (piped) gas service	Good 48	Telephone services