

Exchange Rates and Fundamentals: Footloose or Evolving Relationship?

Lucio Sarno^{a,b} and Giorgio Valente^{c*}

a: University of Warwick

b: Centre for Economic Policy Research (CEPR)

c: Chinese University of Hong Kong

This version: December 2005

Abstract

Using novel real-time data on a broad set of economic fundamentals for five major US dollar exchange rates over the recent float, we employ a predictive procedure that allows the relationship between exchange rates and fundamentals to evolve over time in a very general fashion. Our key findings are that: (i) the well-documented weak out-of-sample predictive ability of exchange rate models may be caused by poor performance of model-selection criteria, rather than lack of information content in the fundamentals; (ii) the difficulty of selecting the best predictive model is largely due to frequent shifts in the set of fundamentals driving exchange rates, reflecting swings in market expectations over time. However, the strength of the link between exchange rates and fundamentals is different across currencies.

JEL classification: F31.

Keywords: exchange rates; economic fundamentals; forecasting.

***Acknowledgments:** This work was partly undertaken while Lucio Sarno was a Visiting Scholar at the International Monetary Fund. The authors are grateful to Philippe Bacchetta, Menzie Chinn, Luca Dedola, Martin Evans, Linda Goldberg, Otmar Issing, Robert Lafrance, Nelson Mark, John Murray, Chris Neely, Mark Taylor, Eric van Wincoop, and to other participants to the 2005 ECB-Bank of Canada Workshop on “Exchange Rate Determinants and Economic Impacts” for comments or useful conversations, and to John Rogers for advice on constructing our real time data set. We are also grateful to Raffaella Giacomini, Ken West, and Halbert White for valuable advice on issues related to the forecasting exercises carried out in the paper. The research reported in this paper was funded by a research grant from the Economic and Social Research Council (ESRC Grant No. RES-000-22-0404). Emmanuel Davradakis provided research assistance. The authors alone are responsible for any errors that may remain.

1 Introduction

Forecasting exchange rates is an hazardous occupation. It is a stylized fact that out-of-sample forecasts of exchange rates produced by structural models based on economic fundamentals are no better than those obtained using a naive random walk or no-change model of the nominal exchange rate. These results, first noted in the seminal work of Meese and Rogoff (1983), have been proved to be very robust. Some recent research, using techniques that account for several cumbersome econometric problems, including small sample bias and near-integrated regressors in the predictive regressions, suggests that models based on canonical fundamentals can explain a small amount of the variation in exchange rates (e.g., Mark, 1995; Mark and Sul, 2001). However, others remain skeptical (e.g., Kilian, 1999; Berkowitz and Giorgianni, 2001; Faust, Rogers and Wright, 2003), so that evidence that exchange rate forecasts obtained using fundamentals models are better than forecasts from a naive random walk remains elusive (e.g., Neely and Sarno, 2002; Cheung, Chinn and Pascual, 2003). Moreover, prior research has documented that economic fundamentals do not contain information useful for forecasting exchange rates partly because the behavior of the ‘virtual fundamentals,’ that is the fundamentals that would be capable of explaining exchange rate movements, is radically different from the behavior of the fundamentals suggested by exchange rate determination theory. Notably, observing the marked increase in volatility of exchange rates which occurs when moving from a fixed or pegged exchange rate regime to a floating regime, Baxter and Stockman (1989) and Flood and Rose (1995) argue that any tentatively adequate exchange rate model should have fundamentals which are also much more volatile during floating rate regimes. This evidence may be seen as suggesting that there are speculative forces at work in the foreign exchange market which are not reflected in the usual menu of economic fundamentals. These constitute the key elements of the ‘exchange rate disconnect puzzle’ (Obstfeld and Rogoff, 2000).

Notwithstanding this evidence, it seems hard to believe that broad knowledge of the state of the economy at a point in time is useless information to forecast exchange rate movements. It may be that exchange rate models perform poorly not (only) because the information in the fundamentals is deficient, but because volatile expectations and departures from rationality are likely to account for the failure of exchange rate models. For example, Frankel (1996) argues that exchange rates are detached from fundamentals because of ‘swings’ in expectations about future values of the exchange rate, listing several pieces of evidence suggesting that expectations are to blame for such behavior. In this line of reasoning, recently Bacchetta and van Wincoop (2004) have provided a theory of exchange rate determination which incorporates the fact that practitioners in the foreign exchange market regularly change the weight they attach to different economic variables—as evidenced in a variety of survey studies (e.g. Cheung and Chinn, 2001)—in the context of a stylized rational-expectations model of

exchange rate determination. In this model, as the market rationally searches for an explanation of the observed exchange rate change, it may attribute it to some macroeconomic indicator, which in turn becomes the ‘scapegoat’ and influences trading strategies. This implies that agents attach an excessive weight to the ‘scapegoat’ variable to explain exchange rate movements during some periods, so that the exchange rate is unrelated to other observed economic fundamentals. The model is also capable of rationalizing parameter instability in empirical exchange rate models of the kind often documented in the relevant literature (e.g. Rossi, 2005a,b). Over time different observed variables may be taken as the scapegoat, so that the weights attributed to economic variables change.

Surprisingly little attention has been directed by the empirical literature towards assessing the potential of these considerations for establishing an economically meaningful relationship between exchange rates and fundamentals. The present paper fills this gap. We build an empirical framework that explicitly takes into account the survey and empirical evidence that there are changing weights in fundamentals and that no fundamentals model appears to perform well for long periods of time. Our framework allows for a menu of fundamentals that comprises not only the standard monetary fundamentals most commonly used in the literature since Mark (1995), but also other variables suggested by exchange rate determination theory, including net foreign assets, interest rates, the trade balance, and lagged values of exchange rate changes. We employ a recursive procedure where we select, quarter by quarter, the best model on the basis of a variety of statistical and economic criteria within all possible combinations of fundamentals, allowing for the explanatory variables and the parameters to change over time. We perform an out-of-sample forecasting exercise using a novel real time data set for the fundamentals, for each criterion used in the model selection procedure and for several major dollar exchange rates. The results of this out-of-sample exercise are compared with the results that would have been obtained if one had knowledge of the best-performing fundamentals model. This comparison sheds light on the usefulness of the information embedded in the fundamentals for forecasting the exchange rate as well as on the ability of the model selection procedure to use such information optimally.

We focus on two key questions. First, as a preliminary to the forecasting exercise, we perform a virtual search for the best performing exchange rate model given a broad set of fundamentals available in real time over the recent float for dollar exchange rates. In other words, we ask whether, allowing for shifts in the weights attached to the fundamentals over time, the best model that optimizes such information is indeed capable of predicting the exchange rate with a reasonable degree of accuracy. Second, when the previous exercise suggests that there is a (time-variant) fundamentals model capable of explaining and predicting exchange rate movements, we ask whether it is possible to recover the best model obtained earlier when the out-of-sample forecasting exercise is conducted using conventional

model selection criteria, period by period, over our sample.

The results, obtained using five major US dollar exchange rates during the recent float and considering short-term forecasts of one-quarter ahead, are as follows. First, the information embedded in the economic fundamentals can explain future exchange rate movements with a remarkable degree of accuracy for three out of five exchange rates. However, this requires the ability to select the best model from the various set of models that can be used on the basis of information available, i.e. this requires that the investor must have available a reliable model selection criterion to discriminate among different specifications of the fundamentals model. Second, if conventional (statistical and/or economic) model selection criteria are used to choose *ex ante* the best model from a large pool of models available, the same set of economic fundamentals is not useful in forecasting exchange rates out of sample. This finding appears to be due to the inability of existing model selection criteria to identify the predictive variables to be used *ex ante* and to capture the frequent shifts in the best model capturing the evolving dynamic relationship between exchange rates and fundamentals.

We find that models which optimally use the information in the fundamentals change often and this implies frequent shifts in the parameters. Our interpretation of these results is that standard model selection criteria appear to be unable to capture such shifts, yielding empirical exchange rate models that cannot forecast the exchange rate better than a random walk model. Therefore, while the stylized fact that fundamentals models cannot beat a no-change model is—yet again—confirmed in this paper, the reason for this result is, on the basis of the evidence presented here, different from many other studies. Our evidence suggests that the exchange rate disconnect phenomenon is unlikely to be caused by lack of information in the fundamentals, and more likely to be the result of poor model selection criteria in this context.

The remainder of the paper is organized as follows. In Section 2 we discuss the vast body of research to which this paper is related and further motivate our approach. Section 3 describes the framework used to analyze exchange rate predictability allowing for shifts in fundamentals and parameter instability. Next, in Section 4, we describe our data set and perform a preliminary empirical analysis of the time series of interest. In Section 5 we discuss our empirical results relating to the virtual search for the best exchange rate model, conditioning on the information in our real-time data set but optimizing such information appropriately. In Section 6 we report the results from the out-of-sample forecasting exercise based on the same set of fundamentals information and employing conventional model selection criteria to choose the best fundamentals model period by period. Section 7 provides a discussion of possible extensions of our research and reports some robustness checks. Section 8 concludes. Details of the construction of our real time data set and of the bootstrap procedure employed in the empirical analysis are provided in the Appendix.

2 Exchange Rates and Fundamentals: Overview and Motivation

2.1 The Ingredients of the Exchange Rate Disconnect Puzzle

A large literature has investigated the relationship between the nominal exchange rate and the menu of fundamentals typically suggested by exchange rate determination theory by focusing on the deviation, say u , of the nominal exchange rate from its fundamental value:

$$u_t = s_t - f_t, \tag{1}$$

where s denotes the log-level of the nominal bilateral exchange rate (the domestic price of the foreign currency); f is the long-run equilibrium of the nominal exchange rate determined by economic fundamentals; and t is a time subscript. The fundamentals term is, most commonly, given by a parsimonious set of ‘monetary fundamentals,’ comprising the differential in money supply and the differential in output. Here f may be thought of as a generic representation of the long-run equilibrium exchange rate implied by modern theories of exchange rate determination (Mark and Sul, 2001).¹

However, it is clear that a broader set of fundamentals could be employed on the basis of international macroeconomics theory. Additional fundamentals include, for example, net foreign assets, as suggested by overlapping generations versions of the new open-economy macroeconomics (NOEM) models (e.g. Cavallo and Ghironi, 2002; Obstfeld, 2004), and the trade balance, as suggested by portfolio balance models (e.g. Branson, 1984; Kumhof and Van Nieuwerburgh, 2005) and elasticity models of the balance of trade (e.g. Kreuger, 1983; Rose and Yellen, 1989; Obstfeld and Rogoff, 2004). Furthermore, Andersen, Bollerslev, Diebold and Vega (2003) document that the set of macroeconomic news that affects the conditional mean of the exchange rate is quite broad, including money, income, prices and other macroeconomic variables.

While it has been difficult to establish the empirical significance of the link between monetary fundamentals and the exchange rate due to a number of cumbersome econometric problems (e.g. Mark, 1995; Kilian, 1999; Berkowitz and Giorgianni, 2001), some recent research suggests that macroeconomic fundamentals move together with the nominal exchange rate over long periods of time (Groen, 2000; Mark and Sul, 2001; Rapach and Wohar, 2002; Rossi, 2005a,b). This result implies that future movements in the exchange rate may be predicted using current information on the departure of the

¹For example, this fundamentals terms is implied by the monetary approach to exchange rate determination (Frenkel, 1976; Mussa, 1976, 1979; Frenkel and Johnson, 1978) as well as by Lucas’ (1982) equilibrium model and by several new open economy macroeconomics models (see Obstfeld and Rogoff, 1995, 2000; and the references in the surveys of Lane, 2001). Hence, the link between monetary fundamentals and the nominal exchange rate is consistent with both traditional models of exchange rate determination based on aggregate functions as well as with more recent microfounded open economy models.

current exchange rate from its value consistent with macroeconomic fundamentals. This rationale has been used to investigate long-run predictability regressions of the following form

$$\Delta_k s_{t+k} = c + \psi_k u_t + \varepsilon_{t+k} \quad (2)$$

where Δ_k denotes the k -difference operator. Indeed, equation (2) is the equation analyzed by a vast literature investigating the ability of monetary fundamentals to forecast the nominal exchange rate out of sample at least since Mark (1995).² In some sense, this predictive regression may be seen as a parsimonious specification of a more general model where the change in the exchange rate is a function of lagged values of information embedded in the menu of economic fundamentals suggested by the theory.³

In general, long-horizon regressions have been used extensively in the literature, but with mixed success (see Kilian, 1999; Berkowitz and Giorgianni, 2001; Kilian and Taylor, 2003; Rossi, 2005a,b). Moreover, the literature described in this section often ignores the fact that data for economic fundamentals may not be available at the time forecasts are made or may suffer from measurement errors. Employing data in real time—that is data that would have been available to researchers at the time forecasts would have been produced—further complicates analyses of the predictive power of fundamentals on the exchange rate (Faust, Rogers and Wright, 2003).

Flood and Rose (1995) shed light on the relation (or lack of it) between the exchange rate and fundamentals in what they term ‘a virtual quest for fundamentals.’ The underlying idea is to compare the volatility of the ‘traditional fundamentals’, say the classic set of monetary fundamentals typically employed in the literature, to the volatility of the ‘virtual fundamentals,’ that is the fundamentals that would be capable of explaining the excess volatility of foreign exchange excess returns. Observing the marked increase in volatility of exchange rates which occurs when moving from a fixed exchange rate regime to a floating regime, Flood and Rose argue that any tentatively adequate exchange rate model should have fundamentals which are also much more volatile during floating rate regimes. In their empirical work, they find little change in the volatility of economic fundamentals suggested by flexible-price or sticky-price monetary models across different nominal exchange rate regimes for a number of major exchange rates. Similar evidence is reported by Baxter and Stockman (1989),

²Note that equation (2) implicitly assumes that deviations from long-run equilibrium are restored via movements in the exchange rate. However, it seems possible that they may be restored also via movements in the fundamentals (see Engel and West, 2005).

³In a series of forecasting tests over horizons ranging from 1 to 16 quarters ahead for a number of dollar exchange rates, Mark (1995) finds that a simple regression of the form (2) may be able to predict the nominal exchange rate only at long horizons, such as the 3- and 4-year horizon, but the Meese-Rogoff results that the nominal exchange rate is not forecastable within the year is found to be robust. Mark interprets this as evidence that, while quarter-to-quarter exchange rate movements may be noisy, systematic movements related to the fundamentals become apparent in long-horizon changes. Predictability is also found to be economically important in the context of optimal asset allocation problems where the criterion of forecast evaluation is based on utility calculations (e.g. see Abhyankar, Sarno and Valente, 2005).

who examine the time-series behavior of a number of key macroeconomic aggregates for 49 countries over the postwar period. Although they detect evidence of higher exchange rate variability under flexible exchange rates than under pegged nominal exchange rate regimes, Baxter and Stockman find no systematic differences in the behavior of the macroeconomic aggregates under alternative exchange rate arrangements. Again, this suggests that there are speculative forces at work in the foreign exchange market which are not reflected in the usual menu of macroeconomic fundamentals.

Taking an alternative view, Engel and West (2005) demonstrate that in a rational-expectations present value model, under the assumptions that fundamentals are nonstationary and the discounting factor is near one, the exchange rate will behave as a ‘near’ random walk process. This implies that the difficulty to predict exchange rates using fundamentals may well be consistent with conventional exchange rate determination models. Engel and West also find that evidence that there is predictive ability from exchange rates to fundamentals. Put differently, the Engel-West framework indicates that the kind of (small degree) of predictability of exchange rates reported in the literature should not be too surprising if one is willing to accept the above assumptions.

2.2 Model Instability and Swings in Expectations

A different line of attack to empirical exchange rate models involves employing a time-varying parameter model. A number of authors have argued that the poor forecasting performance recorded in this literature may be due to the fact the parameters in the estimated equations are unstable. This instability may be rationalized on a number of grounds, in response to policy regime changes as an example of a Lucas critique problem (Lucas, 1976), or because of implicit instability in the money demand or PPP equations, or also because of agents’ heterogeneity leading to different responses to macroeconomic developments over time (e.g. see Schinasi and Swamy, 1989; Rossi, 2005a,b).

But that begs the question as to why exchange rate models perform so poorly. Many researchers have claimed that volatile expectations or departures from rationality are likely to account for the failure of exchange rate models. For example, Frankel (1996) argues that exchange rates are detached from fundamentals by swings in expectations about future values of the exchange rate. Several pieces of evidence suggest that expectations are to blame for such behavior: 1) Survey measures of exchange rate expectations are very poor forecasters and the expectations, themselves, are frequently not internally consistent (Frankel and Froot, 1987; Sarno and Taylor, 2001). 2) Failure of rational expectations is often blamed for the failure of uncovered interest parity (Engel, 1996). 3) Trend following trading rules appear to make positive risk-adjusted excess returns, in apparent violation of the efficient markets hypothesis (e.g. Levich and Thomas, 1993; Neely, Weller, and Dittmar, 1997). 4) Switching from a fixed exchange rate to a floating rate—which changes the way expectations are

formed—changes the behavior of nominal and real exchange rates and the ability of UIP to explain exchange rate changes (Baxter and Stockman, 1989; Flood and Rose, 1995).

This latter point requires some explanation. Fixed exchange rates anchor investor sentiments about the future value of a currency because of the government’s commitment to stabilize its value. If expectations are based on fundamentals, rather than based on changing expectations that do not obey the rational expectations paradigm, then the relationship between fundamentals and exchange rates should be the same under a fixed exchange rate regime as it is under a floating regime. This is not the case. Countries that move from floating exchange rates to fixed exchange rates experience a dramatic change in the relationship between prices and exchange rates. Specifically, real exchange rates (exchange rates adjusted for inflation in both countries) are much more volatile under floating exchange rate regimes, where expectations are not tied down by promises of government intervention (Mussa, 1986). This result suggests that swings in investor expectations may detach exchange rates from fundamental values in the short run. Similarly, exchange rate models seem to do such a poor job explaining exchange rates while doing a pretty good job with semi-fixed rates such as those found in the European Monetary System (Flood and Rose, 1996).⁴

In this line of reasoning, recently Bacchetta and van Wincoop (2004, 2005) have provided a theory of exchange rate determination which incorporates the fact that practitioners in the foreign exchange market regularly change the weight they attach to different macroeconomic variables—as evidenced in a variety of survey studies (e.g. Cheung and Chinn, 2001)—in the context of a stylized rational-expectations exchange rate determination model. This model is capable of explaining parameter instability on empirical exchange rate models in terms of a ‘scapegoat’ story, where some variable is given excessive weight during some period, implying movements in the exchange rate that are unrelated with observed economic fundamentals, for example due to unobserved liquidity trades. As the market rationally searches for an explanation for the observed exchange rate change, it may attribute it to some macroeconomic indicator, which in turn becomes the scapegoat and influences trading strategies. Over time different observed variables may be taken as scapegoat, so that the weights attributed to macroeconomic variables change.⁵

⁴See also Flood and Rose (1999), who develop a canonical model of the exchange rate that explains why uncovered interest parity—and exchange rate forecasts—might perform poorly in the short term even with perfect rational agents.

⁵The key feature of the model is, therefore, the heterogeneity of views in the market about the importance of macro variables in explaining the exchange rate and the disagreement about the true course of observed exchange rate fluctuations. The heterogeneity is modelled as private signals about structural parameters. Investors do not know whether these fluctuations are due to unobserved fundamentals, say liquidity trades, or to a larger than expected weight on certain observed macro fundamentals. In this environment, agents naturally blame the variables they can observe, that is the fundamentals. See also Bacchetta and van Wincoop (2004, 2005).

2.3 Introducing Our Approach

In this paper, we build an empirical framework that explicitly takes into account the survey evidence that there are changing weights in fundamentals and that no model of fundamentals appears to perform well for long periods of time. Our framework allows for a menu of fundamentals that comprises not only the standard monetary fundamentals most commonly used in the literature since Mark (1995), but also other variables suggested by exchange rate determination theory, including net foreign assets, interest rates, the trade balance, and lagged values of exchange rate changes. We employ a recursive model selection procedure where we select, quarter by quarter, the best model on the basis of a variety of statistical and economic criteria within all possible combinations of fundamentals, allowing for the model specification and the parameters to change. We perform an out-of-sample forecasting exercise using real time data for the fundamentals, for each criterion used in model selection and for several major dollar exchange rates, and then compare the results of this purely out-of-sample exercise with the results that would have been obtained if one knew the best performing model. This comparison will shed light on the usefulness of the information embedded in the fundamentals for forecasting the exchange rate as well as on the ability of the model selection procedure available to us to use such information optimally. In the next section we describe in detail the empirical framework.

3 Methodology

3.1 The Pool of Models

Let us consider a US investor who wishes to forecast in real time a bilateral dollar nominal exchange rate and believes that the exchange rate can be predicted by a menu of economic and financial variables that are suggested by exchange rate determination theory. Since the investor does not know the true data generating process (DGP) linking such fundamentals to the future exchange rate, a reasonable way to produce the exchange rate forecasts then involves searching for an adequate model specification among all possible models believed to be capable of predicting the exchange rate. The search ought to be general enough to encompass all possible structural models of exchange rate determination suggested by the theory and considered plausible by the investor as well as to allow also for the possibility that none of the theoretical specification is indeed useful for forecasting, so that the framework ought to allow for the possibility that the exchange rate follows a random walk (with or without a drift). The decision of the best exchange rate model to use is based on one or more statistical or economic criteria. As the investor repeats this exercise over time and new information (observations) become available, the investor may change the model following the model selection criteria employed. These changes in the model may be due to the learning process of the investor

with the arrival of new information and/or to the changing nature of the DGP of the exchange rate.

In order to characterize this behavior econometrically, we follow the stock price predictive approach of Pesaran and Timmermann (1995, 2000) and Bossaerts and Hillion (1999), which we adapt to the context of exchange rate forecasting. At each point in time t , the investor estimates all possible combinations of $k-1$ fundamentals which may predict future movements in exchange rates, in addition to a drift term in exchange rate changes as a possible regressor. Thus, the full set of possible models within which the investor must select one comprises 2^k different models. The simplest model is a random walk without drift (no regressors) and the richest model is a fully unrestricted model where all k regressors, inclusive of the drift term, are considered. Denote a model M_i for $i = 1, \dots, 2^k$, and let the set of regressors be $X_j = \mathbf{1}, x_1, \dots, x_{k-1}$. Consider a $k \times 1$ selection vector, say v_i , composed of ones and zeros where a one in its j -th element indicates that the regression in question is included in the model. Then, M_i may be represented by the k -digit string of zeros and ones corresponding to the binary code of its number. If k_i is the number of regressors included in model M_i , $k_i = e'v_i$, where e' is a $k \times 1$ vector of ones. Then, M_i may be written as follows:

$$\Delta s_{t+1} = \beta_i' X_{t,i} + \varepsilon_{\tau+1,i} \quad (3)$$

where $\Delta s_{t+1} = s_{t+1} - s_t$, β is a vector of parameters; $X_{t,i}$ is a $k_i \times 1$ vector of regressors (available to the investor at time t) under model M_i ; and $t = 1, \dots, T$. Estimation of all candidate models may be carried out by ordinary least squares (OLS)—see Pesaran and Timmermann (1995, 2000). Applying a model selection procedure (based on statistical and/or economic criteria), the investor will conduct the first estimation of all 2^k models on the basis of the first $\tau < T$ observations and select a certain model. Then, the investor will re-estimate the model period by period as new data become available and repeat the procedure up to time T , yielding time-varying estimates of β_i throughout the sample. In essence, starting from $t = \tau$, the procedure delivers at each point in time estimation of 2^k models with corresponding 2^k one-period-ahead forecasts of the change in the exchange rate. Out of the 2^k models estimated, at each point in time $t = \tau, \tau + 1, \dots, T - 1$, only one is selected as a result of a selection procedure based upon a set of conventional model selection criteria.

3.2 Model Selection

We employ a number of criteria in order to select the forecasting model, within the set of possible models obtaining from all combinations of the fundamentals considered, on the basis of in sample information: the Mean Absolute Error (MAE); Root Mean Square Error (RMSE); the Last Absolute Error

(LAE)⁶; the adjusted coefficient of determination⁷ (\bar{R}^2); the Akaike Information Criterion (AIC); the Schwarz Information Criterion (a Bayesian information criterion, say BIC); a sign criterion (SIGN) based on directional accuracy of the in sample predictions of the exchange rate change; the Posterior Information Criterion (PIC), due to Phillips and Ploberger (1996); and the Fisher Information Criterion (FIC), due to Wei (1992).

3.3 Forecast Evaluation

In terms of criteria of evaluation of out-of-sample forecast accuracy, we again rely both on statistical and economic criteria. The statistical criteria include the MAE and the RMSE. We also utilize two economic criteria of forecast evaluation. First, we use the recursive end-of-period wealth criterion (EOPW), which maximizes the cumulated wealth using forecasts from model M_i in a switching portfolio

$$W_{t,i} = W_0 \prod_{t=1}^{\tau} \left\{ 1 + \left[r_t + \left| \widehat{\Delta}s_{t+1,i} - (r_t - r_t^*) \right| \right] [1 - TC] \right\} \quad (4)$$

where W_0 is initial wealth, and $W_{t,i}$ denotes wealth obtained under Model M_i at the end of the forecasting period calculated using one-step-ahead out-of-sample prediction of the exchange rate changes ($\widehat{\Delta}s_{t+1,i}$) and the corresponding one-period interest rates at home (r_t) and abroad (r_t^*); the hat denotes out-of-sample predictions; the asterisk denotes a foreign variable; and τ is as defined in Section 3.1. The term in absolute value is the predicted foreign exchange excess return for the dollar exchange rate in question, and the absolute value reflects the fact that we allow for short selling and hence the sign is unimportant in determining the investment decision; TC denotes transactions costs.⁸ The second economic criterion we employ is the recursive excess return criterion (RR), calculated as the ratio of the mean portfolio returns to their standard deviations realized out-of-sample:

$$RR_{t,i} = \frac{\frac{1}{\tau} \sum_{t=1}^{\tau} \left[\left| \widehat{\Delta}s_{t+1,i} - (r_t - r_t^*) \right| \right] [1 - TC]}{\sigma_R} \quad (5)$$

⁶The LAE is simply the criterion which selects the model that delivers the smallest in sample prediction error in the last observation available in real time.

⁷The \bar{R}^2 is obtained as the \bar{R}^2 from a regression of the actual data for exchange rate changes on the forecast exchange rate changes over the forecast period, consistent with the regression suggested in this context by West and McCracken (1998).

⁸Equation (4) is consistent with the following one-period wealth calculation at any point in time $t = 1, \dots, \tau$: $W_{t+1} = W_t \{1 + R_{p,t}\}$, where $R_{p,t} = \left[(1 - \omega) r_t + \omega \left(r_t^* + \widehat{\Delta}s_{t+1,i} \right) \right]$ and ω denotes the weight within the portfolio associated with the investment in the foreign asset. In our empirical exercise we assume that at any point in time the investor chooses to invest in a portfolio including only domestic or foreign assets according to the sign of the expected excess return in the foreign exchange market $\widehat{\Delta}s_{t+1,i} - (r_t - r_t^*)$. This, together with the allowance for short-selling, implies that, under the two cases, the value of the weight ω in $R_{t,p}$ is equal to 2 (i.e. short sell the domestic asset and invest the proceedings in the foreign asset) or -1 (i.e. short sell the foreign asset and invest in the domestic asset) respectively. Further, as showed in equation (4), we have explicitly taken into account transaction costs by deducting from the gross return $R_{t,p}$ a fixed amount $TC = 0.0004$ only when there is a shift in the portfolio allocation, or equivalently when the signs of the excess returns in the foreign exchange market at time t and $t - 1$ are discordant.

where σ_R is the standard deviation of the portfolio return calculated over the forecasting period at the net of transactions costs. This criterion is similar to the Sharpe ratio criterion used by Pesaran and Timmermann (1995, 2000).

The primary aim is to select a forecasting equation that could be viewed as being a reasonable approximation to the DGP at the time when this procedure would have been used, making the exercise as realistic as possible. Because our first estimation is done with quarterly data up to 1985, when we produce the first one-quarter-ahead forecast, a US investor would have been able to employ the procedure described in this section as OLS was obviously known and easy to employ at that point in time. Also all of the criteria employed were known at that time, with the exception of the PIC and the FIC. In addition, as discussed in detail below, we base our analysis on a suitably constructed set of real time data which correspond exactly to the data available at the time forecasts were made for each of the economic and financial indicators employed.

3.4 Data Snooping

In evaluating the forecasting models in our setup, it is important to be aware of data-snooping biases. In our framework, where there are 2^k models estimated and evaluated at each point in time, the effects of data snooping can be dramatic. In order to intuitively explain the danger of data snooping biases in our context, let us use an example.⁹ Suppose that we are trying to predict the outcome of a fair coin toss (i.e. we are trying to predict something unpredictable). Let the coin to be predicted be called coin A. The two competing predictors are two other fair coins, say coins B and C. All three coin flips are independent. Then using our methodology and analyzing forecasts *ex post*, we would find that we can predict the outcome of a flip in coin A 75 percent of the time.¹⁰ Naturally, however, if one performed “real time” forecasting, it would not be possible to beat a 50 percent success rate. Therefore, evaluation of forecasting performance when forecasts are obtained from a large pool of models from which a single model is selected *ex ante* requires care in allowing for these biases.

We employ the Reality Check bootstrap procedure proposed by White (2000) for testing the null hypothesis that the best model selected in a specification search has no predictive superiority over a given benchmark model—in our case the random walk model. This procedure is specifically designed to allow aggressive model searching to be undertaken with confidence that one will not mistake results that could have been generated by chance for genuinely good results, effectively correcting for data snooping biases in forecast evaluation—see Appendix B for a description of the procedure.

⁹This example is due to Ken West, who used it in private correspondence and kindly allowed us to report it here.

¹⁰This is because the only time we won't have a correct prediction from at least one of coins B and C is when A=heads and B=C=tails, or A=tails and B=C=heads. There are 8 possible configurations of coin flip results for the three coins, so those 2 events occur with probability 2/8; the implication is that the model search works 6/8=75 percent of the time.

4 Data and Preliminary Analysis

Our data set comprises quarterly observations on money supply, income (gross domestic product), trade balance, 3-month eurorates for the US, Japan, UK, Canada, Switzerland and Germany and spot exchange rates for the US dollar *vis-à-vis* the Japanese yen, UK sterling, Canadian dollar, Swiss franc and German mark. We also have data on net foreign assets between the US and in each Japan, UK, Canada, Switzerland and Germany. The sample period covers most of the recent floating exchange rate regime, from 1977Q1 to 2003Q3 (except for Germany, which ends in 1998Q4), and the start date of the sample was dictated by data availability (for the net foreign assets). Our choice of countries reflects our intention to examine exchange rate data for a set of major industrialized economies that have been governed by a pure float over the sample.¹¹

For the purposes of our analysis we assembled and compiled real time data of quarterly frequency. In particular, we constructed four vintages for each year considered with the first real time observation starting at 1987Q3. We performed the first estimation of the set of models examined using data from 1977Q1 to 1987Q3 and hence the first exchange rate forecast produced refers to the change in the exchange rate from 1987Q3 to 1987Q3. A detailed description of the construction of our data set is given in Appendix A.¹²

Some of the data were transformed in natural logarithms prior to beginning the empirical analysis to yield time series for the nominal exchange rate, s_t , US money m_t , foreign money, m_t^* , US output, y_t , foreign output, y_t^* . We did not take logs for the domestic and foreign interest rates, r and r^* , and the domestic and foreign trade balance, tb and tb^* . Our measure of NFA is calculated as the difference between the log-detrended purchases and sales of foreign assets, consistent with the definition of Cavallo and Ghironi (2002, p. 1074).

The regressors we consider in employing the procedure described in Section 2 are seven. 1) The once-lagged nominal exchange rate change, in order to allow for any potential slight persistence in exchange rate movements (Backus, Gregory and Telmer, 1993). 2) The deviation from a canonical monetary fundamentals model. This deviation we term mf_t , and we construct as $mf_t = (m_t - m_t^*) - (y_t - y_t^*)$ (Mark, 1995). 3) The interest rate differential, $(r_t - r_t^*)$, in the spirit of the conventional

¹¹Note that, while Canada, Japan and Switzerland have experienced a free float since the collapse of the Bretton Woods system in the early 1970s, the UK was in the Exchange Rate Mechanism (ERM) of the European Monetary System (EMS) for about two years in the early 1990s. However, given the short length of this period, we consider sterling as a freely floating exchange rate in this paper. The only mixed-regime country during this sample is Germany, which was in the ERM for most of the sample period under investigation and in fact joined the European Monetary Union on 1 January 1999, when the euro replaced the German mark as Germany's national currency.

¹²Faust, Rogers and Wright (2003) first employed a real-time data set in exchange rate forecasting, focusing primarily on long-horizon regressions based on monetary fundamentals alone. Our real-time data set is broader both in terms of variables and in terms of exchange rates examined, and for some of the variables and vintages used here we had no prior study to follow. We are indebted to John Rogers for guidance and advice on the construction of some of the real time data.

uncovered interest rate parity condition. 4) The net foreign asset position of the US relative to the foreign country in question, NFA_t 5) The US trade balance, tb , 6) The foreign country's trade balance, tb^* and 7) An intercept term. All variables (3)-(6) are lagged once in the empirical investigation discussed below.

As a preliminary exercise, we test for unit root behavior of each of the time series to be used in the predictive framework described in Section 3. Specifically, we test for a unit root in Δs_t , mf_t , $(r_t - r_t^*)$, tb , tb^* and NFA_t over the full sample. We employ the two efficient unit root tests (MZ_α and MZ_t) proposed by Ng and Perron (2001); these tests use generalized least squares-detrending to maximize test power and a modified information criterion to select the lag truncation in order to minimize size distortion. The results—not reported to conserve space but available upon request—suggested for each of these six time series and for each country examined, the unit root hypothesis is rejected at conventional nominal significance levels. In turn these results have a clear economic interpretation, confirming that the change in the exchange rate is stationary, the level of the exchange rate co-moves with standard monetary fundamentals over the long run, the interest rate differential across these major economies is stationary, and trade balances and net foreign asset positions also revert to some long-run equilibrium level.

5 Empirical Results I: A Virtual Search for the Best Fundamentals Model

5.1 Can We Beat the Random Walk with a Fundamentals Model?

In our empirical work, we use the procedure described in Section 3 to shed light on several issues. The first question we address is whether fundamentals do contain information that is useful in predicting exchange rate movements. The digression in Section 2 on the related literature on swings in expectations in the foreign exchange market and theoretical contributions based on the ‘scapegoat’ view of exchange rate movements as well as survey evidence on the behavior of agents in foreign exchange market suggests that the relationship between exchange rate movements and the fundamentals is likely to be characterized by some fundamentals having much more impact than others for certain periods of time and by severe parameter instability. Our predictive framework allows both for the possibility that the set of fundamentals determining the exchange rate changes over time and for the possibility of the parameters to change period by period in a recursive fashion. Hence, our procedure ought to be able to shed light on the role of these phenomena in understanding the link between exchange rates and fundamentals. The only caveat is that one needs reliable model selection criteria that are capable of selecting, within the broad set of fundamentals considered in the information set, the best empirical

exchange rate model. To be immune from this problem, before carrying out a pure out-of-sample forecasting exercise in real time where models are selected with specific criteria, we carry out a simpler exercise. Specifically, we ask whether the best model (BM) constructed at a point in time t , on the basis of information on the fundamentals available at time t , is a model consistent with the exchange rate disconnect results. In order to carry out this exercise, we implemented the predictive procedure described in Section 3 and chose BM to be the model that minimizes the absolute error between the ex post realization of the exchange rate change and the model forecast exchange rate change, i.e. we selected BM as the model M_i that

$$\min_i \left| \Delta s_{t+1} - \widehat{\Delta s}_{t+1,i} \right| \quad t = \tau, \tau + 1, \dots, T - 1 \quad (6)$$

where $\widehat{\Delta s}_{t+1,i}$ is the one-quarter-ahead out-of-sample forecast produced by model M_i . BM, constructed according to this rule, is the model which has performed better in forecasting the one-quarter-ahead exchange rate and is, therefore, the model that would have been selected if one had perfect knowledge on what is the best combination of the seven regressors in the predictive framework at each point in time.

Note that this is different from assuming perfect foresight in that the BM is a model made of fundamentals information—i.e. the exchange rate forecasts $\widehat{\Delta s}_{t+1,i}$ are generated from each model M_i conditional only on information available at time t . If such information is of no value in predicting the exchange rate, or if the exchange rate follows a random walk, the BM will still be unable to outperform a random walk benchmark after accounting for data snooping bias.¹³ We carry out this exercise for each dollar exchange rate considered, and compare the results from the BM obtained from our set of fundamentals to the scenario that one would obtain from using a simple random walk (RW). This scenario is of course ‘virtual,’ it is a strong abstraction from reality. However, this analysis is useful in establishing how well (badly) one can forecast the exchange rate using economic and financial fundamentals within our framework relative to the case where the exchange rate cannot be predicted. For each scenario, we calculate the MAE and the RMSE. If it is not possible to forecast exchange rates (better than a random walk model) with such virtual search, then there would be no point implementing the procedure in Section 2 since we would already know that the outcome would be a failure of the broad set of fundamentals to forecast exchange rates.

The results from this exercise are presented in Table 1. The MAE and RMSE for the BM is lower than the corresponding MAE and RMSE from a random walk model. However, the improvement in terms of MAE and RMSE is statistically significant only for the Japanese yen and the UK sterling at the 5 percent significance level, and at the 10 percent level for the Swiss franc, when p -values

¹³This is also different from an in-sample evaluation since the models are used to forecast the exchange rate conditional on information available at the time of the forecasts when the criterion in (6) is applied.

calculated by bootstrap using White’s Reality Check procedure to account for data snooping biases (White, 2000). Therefore, on the one hand, for three major dollar exchange rates, our exercise suggests results, at first glance, in contrast with the view that the information embedded in economic and financial fundamentals have no predictive power for the exchange rate. On the other hand, for two exchange rates, namely the Canadian dollar and the German mark, we find that the random walk simply cannot be outperformed by our recursive procedure, even with the huge informational advantage used in our procedure. Given this, it would be impossible to outperform the random walk when doing the exercise *ex ante*, i.e. without such informational advantage. From this point onwards, therefore, we do not attempt to forecast the Canadian dollar and the German mark and accept the fact that our broad set of fundamentals will not be capable of producing a model that is better than a random walk. Instead, we shall focus entirely on the analysis of the remaining three exchange rates.¹⁴

The accuracy of the forecasts achieved from the BM for the dollar against the Japanese yen, UK sterling and Swiss franc is made visually clear by the graphs in Figure 1, which presents the actual exchange rate change over the forecast horizon and the forecast produced by the BM for each of these exchange rates. The graphs make apparent how the exchange rate forecasts from the BM track the exchange rate closely even though they display lower variability. This lower variability can be attributed to either noise that is unrelated to fundamentals or to a missing fundamental; the figure confirms the stylized fact that the exchange rate is excessively volatile relative to the fundamentals. However, such excess volatility is much lower than what is typically recorded in the relevant literature and does not prevent the BM fundamentals model from predicting fairly accurately exchange rate changes and more accurately than a random walk benchmark, even when allowing carefully for data snooping.

5.2 A Simple Check on Simulated Data

In order to re-assure ourselves that the superiority to a random walk specification was genuine, we also employed the virtual search replacing the observed data for the exchange rate with artificial data simulated under a driftless random walk. Using these artificial data, if the Reality Check works well, one should not find any evidence of superior predictability of the BM relative to the random walk benchmark. In fact, using these simulated data we recorded Reality Check *p*-values in the range from 0.97 and 0.99 for the exchange rates examined. This is a comforting robustness check on the validity

¹⁴While for the German mark some results in favor of predictability exist in the literature (e.g. Mark, 1995), with respect to the Canadian dollar our results are not surprising. It is well documented that conventional fundamentals do not play an important role in explaining short-term fluctuations in the Canadian dollar and that a key short-run explanatory variable is commodity prices, reflecting the nature of this exchange rate as a classic commodity currency. E.g. see Amano and van Norden (1995) and Chen and Rogoff (2003).

of the Reality Check procedure implemented in our specific context, which raises confidence on the genuinity of the superior performance of the BM in the case of the Japanese yen, UK sterling and Swiss franc.

5.3 How Frequently Does the Best Exchange Rate Model Change?

A logical question to ask relates to the implications of these results for the identification of the true DGP governing the dynamics of these three exchange rate. To address this issue we calculated the number of times the procedure outlined above selected the random walk (with or without drift) as the BM. By looking at this number we find that, over the forecasting period and for all exchange rates examined, the random walk model was selected at most once; and in one case, the UK, the random walk model was never selected as the BM. This finding suggests that the true DGP of the exchange rate is not represented by a pure random walk and there are large economic and statistical gains from appropriately combining economic fundamentals over time in order to exploit their information content to predict out-of-sample the exchange rate.

Given the apparent ability to produce accurate exchange rate forecasts under the BM scenario, one then wonders which fundamentals are driving this result and how often they are switching. Figure 2 (left panel) plots the ‘inclusion frequency’ (IF) for each of the regressors for the dollar-yen exchange rate, which we use as a representative exchange rate. The IF is a binary time series which indicates when a regressor is included in the BM, taking the value of unity, and when it is not included, taking the value of zero. The graphs of the IF for each of the seven regressors suggest that all of them have some importance in the BM, but they are hardly ever important for a long time. In other words, the model using the information set available to the investor (forecaster) optimally changes quite often, as suggested by the high number of switches that characterize each of the explanatory variables in the regression. This seems consistent both the frequent swings in expectations that Frankel (1996) blames for the difficulty to explain exchange rate movements with standard fundamentals models, as well as with the story about higher order beliefs that is behind the scapegoat model of Bacchetta and van Wincoop (2004) and their conjecture that agents in the foreign exchange market often change the weight they attach to a fundamental and their view on whether a fundamental matters or not for determining the future value of the exchange rate. Investigation of the persistence of the inclusion frequencies shows that on average, across macroeconomic fundamentals employed and exchange rates investigated, the variables in the information set are selected for about two quarters. However, this persistence exhibits a very diverse behavior when looking across exchange rates and at different variables one at the time. In the case of the Japanese yen, for example, the highest persistence is recorded by the inclusion frequency of the deviation from monetary fundamentals (3 quarters), followed

by the net foreign asset position and the foreign trade balance (2.4 and 2.3 quarters respectively). The least persistent variables are the interest rate differential and the US trade balance which exhibit an average inclusion of 1.9 and 1.3 quarters respectively.

Overall, across the three exchange rates for which the BM beats a random walk, we can identify two groups of variables: one, which we may label short-term fundamentals, that has a frequent impact on the exchange rate but only for a very short periods of time (i.e. one or two quarters). In this group we find the net foreign asset position, the US trade balance and, in most cases, the interest rate differentials. There is another group of variables (say, long-term fundamentals) whose switches are less frequent but more persistent (4 quarters or more). In this group we find the deviations from monetary fundamentals and the foreign trade balance. This simple descriptive evidence suggests that investors revise their expectations at discrete intervals, generating over time frequent switches in the predictive regressions between fundamentals.

Given that switches for some fundamental variables appear to occur over time at a relative high pace, it is interesting to understand if they exhibit persistence of some kind or are independently distributed. This is important because, if the shifts are independently distributed, it would be impossible to predict them. To address this issue we used a fairly general test for the null hypothesis that the switches are independent and identically distributed against an unspecified form of dependence, namely the Brock, Dechert, Scheinkman and LeBaron (1996) test (BDSL). The results, reported in Table 2, show that for all exchange rates examined, the null hypothesis of independence of all switches is strongly rejected with p -values virtually equal to zero. This evidence is also generally confirmed for individual variables, where the null hypothesis of independence is rejected in the majority of cases, as well as for all variables jointly.¹⁵

5.4 Summing Up

To summarize, the results in this section suggest that the information in the fundamentals, when used in a framework that allows for the set of fundamentals in the model change over time, can forecast some exchange rates with a reasonable degree of accuracy. However, this requires the ability to select the best model from the various set of models that can be used on the basis of information available, i.e. this requires that the investor must have a reliable model selection criterion to discriminate among different specifications of the fundamentals model. The selection criterion would have to be able to generate a high, but not random, number of switches in the model since the BM scenario clearly indicates that the weight attached to a fundamental can vary substantially over time.

¹⁵The p -value for the null hypothesis of joint independence is calculated by using Bonferroni bounds. In particular, for arbitrary correlations among BDSL statistics, the upper bound p -value for the joint test is given by $p \leq \min(k \cdot \min(p_1, \dots, p_k), 1)$ providing the strongest evidence against the null hypothesis (see Kounias, 1968; Hunter, 1976; Galambos and Simonelli, 1996; Sullivan, Timmermann and White, 1999 and the references therein).

6 Empirical Results II: Real-Time Selection of the Best Fundamentals Model

6.1 Can We Select the Best Exchange Rate Model in Real Time?

We now turn to the empirical implementation of the procedure described in Section 3, applied using data in real time and, unlike the previous section, employing the model selection criteria described in Section 3. Given the results from the ‘virtual’ out-of-sample exercise carried out in the previous section, if the model selection criteria perform satisfactorily, an investor forecasting three major dollar exchange rates should have been able to perform rather well over the sample 1985-2003.

The results from implementing the recursive predictive procedure provide very different evidence. In Tables 3 and 4 we report the out-of-sample results, for each model selection criterion considered. As for the virtual exercise, we report in Table 3 the MAE and the RMSE while in Table 4 we report for each exchange rate investigated the economic criteria of evaluation described in Section 3 (i.e. EOPW and RR). The results indicate that none of the model selection criteria can replicate, even remotely, the best model that uses optimally the information in the fundamentals available in the information set of the investor. The differences in MAEs and RMSEs relative to a random walk are miniscule, and the other two economic criteria (EOPW and RR) deliver similar unsatisfactory outcomes.

While the comparison of the recursive predictive procedure in real time with the best model may be seen as unfair given the huge informational advantage offered to the virtual search carried out in Section 5, the comparison with the random walk model is a reasonable one, given that the random model is the natural benchmark in the relevant literature. The latter comparison reveals that, in general, the performance of the random walk model (which allows for a drift term estimated recursively) is very similar to the performance of the recursive predictive procedure (irrespective of the model selection criterion) for all exchange rates examined. We formally test the null hypothesis that the recursive procedure and the random walk model have equal forecast accuracy by employing the White’s (2000) Reality Check procedure and report the p -value for this test in parentheses. The p -values are, for each exchange rate and selection criterion, comfortably larger than conventional nominal significance levels, suggesting that the recursive predictive procedure based on our set of fundamentals is, apart from few scattered exceptions, indistinguishable in terms of out-of-sample performance from a random walk model that uses no information about the state of economy. In short, we confirm the stylized fact in the empirical literature that fundamentals models cannot beat a random walk model in short-term exchange rate forecasting. In fact, we refine this stylized fact and prove its robustness even in the context of a recursive procedure that allows for a fairly rich set of fundamentals and allows the model specification (fundamentals) to change period by period.

6.2 Why Can't We Use the Information in the Fundamentals?

These results beg the question: why is there such a striking contrast between the results in the 'virtual' out-of-sample exercise (Table 1) and the results in the recursive real-time procedure. The most logical answer must be, at least partly, that the model selection criteria employed cannot select the best model we produced in the virtual exercise since the virtual analysis in Section 5 differs from the analysis in Section 6 only due to the model selection criterion (i.e. the pool of models and forecasts are the same in these two exercises).

In order to shed light on the key differences between the model selected using model selection criteria and the best model generated in Section 5, we present in Figure 2 (right panel) plots of the 'inclusion frequency' (IF) for each of the regressors for the dollar-yen exchange rate and for the case where the \bar{R}^2 is the model selection criterion, which may be seen as a representative criterion. The resulting IFs—again a binary time series which indicates when a regressor is included in the selected model, taking the value of unity, and when it is not included, taking the value of zero—for each of the seven regressors suggest that all of them except the lagged exchange rate are included in the model over the forecast sample according to the model selection criterion. However, comparison with the corresponding IFs from the best model given on the left panel of Figure 2 reveals simply too clearly that fundamentals tend to be selected for a long time. In other words, the model using the information set available to the investor (forecaster) according to the model selection criterion tends to keep a fundamental for many years, inducing huge persistence in the IFs. The number of switches is much smaller than the number of switches recorded under the best model. This is especially true for the interest rate differential and net foreign assets, which are included in the model throughout the forecast period, and also for the drift term, which is included ever since 1988. The massive persistence of the IFs suggests that using conventional model selection criteria it is not possible to reproduce the frequent swings that, according to our 'virtual' out-of-sample exercise, characterize the sort of the model that can explain future movements in the exchange rate. The results obtained using standard criteria would be in contrast with the existence of frequent shifts in expectations advocated by Frankel (1996), or the scapegoat story of exchange rate determination of Bacchetta and van Wincoop (2004), as well as with the survey evidence that agents in the foreign exchange market *often* change the weight they attach to a fundamental (Cheung and Chinn, 2001).

Figure 3 further highlights the properties of the forecasts generated under the recursive procedure in real time and the simple random walk model, together with the forecasts generated by standard monetary-fundamentals model of the kind used by Mark (1995). The graphs suggest that the random walk forecasts are the most persistent, followed by the forecasts of the monetary-fundamentals model and the more volatile forecasts implied our procedure. However, it is also apparent how our procedure

generates exchange rate forecasts which cannot match many of the turning points in the exchange rate change and, in fact, occasionally make persistent errors.¹⁶

6.3 Summing Up

Overall, taking together the results in Sections 5 and 6, it is clear that the information embedded in the conventional menu of fundamentals may be useful in forecasting exchange rates out of sample. In fact, the forecasts that can be produced by selecting the best set of fundamentals at each point in time can be very satisfactory on the basis of conventional statistical and economic criteria of forecast evaluation. However, the model that optimally uses the information in the fundamentals changes often and is consistent with frequent shifts in the parameters. Standard real-time model selection criteria appear to be unable to generate such shifts, yielding empirical exchange rate models that cannot forecast the exchange rate better than a random walk model. Therefore, while the stylized facts that fundamentals models cannot beat a no-change model is confirmed in this paper, the reason for this result is, on the basis of the evidence presented here, different from many other studies. Our evidence suggests that the exchange rate disconnect puzzle arising from our result is unlikely to be caused by lack of information in the fundamentals, and more likely to be the result of poor model selection criteria in this context, presumably because of the instability that characterizes the unknown true process driving exchange rates.

7 Robustness

In this section we report some of the robustness checks we carried out to assess the sensitivity of the results reported in Sections 5-6. First, we repeat the exercises carried out in Sections 5-6 using historical data for all fundamentals that are not in real time, i.e. using the revised data available at the present time but not at the time the investor was making the forecasts. Faust, Rogers and Wright (2003) first examined the importance of using data in real time in the context of long-horizon predictability regressions of the type used by Mark (1995). Using dozens of different data vintages for money and output, Faust, Rogers and Wright (2003) show that the predictability found in Mark's study tends to be present only during a two-year window of data around the time of that study. Predictability appears to be weaker during other time periods. However, Faust, Rogers and Wright also show that models using data in real time perform better in out-of-sample forecasting

¹⁶As for the 'virtual' model selection, we have carried out a similar exercise to understand how many times each individual selection criterion would have selected the random walk model (with or without drift) as the preferred model. In general most of the criteria, over the forecasting period and for all exchange rate examined, have never selected the random walk model as preferred model. However, interesting exceptions are represented by the BIC and RR criteria which on average have selected the random walk model 63 and 15 percent of the times respectively.

than models using revised data. In Table 5, we report the results for the dollar-yen exchange rate, as a representative case. Panel A gives the results for the ‘virtual’ exercise and hence these results should be compared to the corresponding results in Table 1. Such comparison reveals that using revised has no impact on the results for the best model. However, the results when the forecasting model is selected using standard statistical criteria (Panel B of Table 4) reveals that using revised data generally tends to produce similar, albeit slightly lower, values of the \overline{R}^2 out of sample. In terms of economic forecast evaluation criteria, it is difficult to decide whether the results using revised data are better or worse than the results using vintages, since the outcomes is different depending the selection criterion. However, the differences are quantitatively small.

A second robustness check involves using a rolling, rather than recursive, procedure for forecasting. One may argue that, if the recursive procedure employed in the core of the paper fails to forecast exchange rates because of its inability to generate enough switches in the forecasting model, then this problem may be mitigated by employing a rolling scheme for forecasting since this relies less on the distant past. We therefore repeated the empirical analysis using a rolling forecasting scheme with a rolling window of 5 years (i.e. 20 quarterly observations). On the other hand, the rolling scheme is more subject to a constant estimation risk (parameter uncertainty), whereas in the recursive procedure the estimation risk reduces as the number of observations increases. The results, given in Table 6 using again the dollar-yen as representative exchange rate, suggest that there are negligible differences in the case for the ‘virtual’ out-of-sample exercise (Panel A), while for the real-time exercise (Panel B) the rolling forecast scheme tends, in general, to perform worse than the recursive scheme both in terms of statistical and economic forecast evaluation criteria. These differences are, however, very marginal.

A third robustness exercise we present involves adding a second lag in each of the regressors used in the information set. In order to avoid a drastic increase in the number of models to be estimated, we assume that, although the parameters on the first and second lag for a fundamental are different, they switch together—i.e. switches remain associated with each fundamental, and each fundamental is now identified as two lags of the variable in question. The results from this exercise, reported again for the dollar-yen exchange rate (Table 7), suggest that the addition of a second lag induces at best marginal improvements in the forecasting performance.

Overall, the robustness exercise reported in this section indicate that the key results reached in Section 5-6 are fairly robust: the information in the fundamentals can, in theory, be used optimally to forecast the exchange rate with a high degree of accuracy; a recursive procedure where the forecasting model is chosen using conventional model selection criteria cannot select the best exchange rate model and cannot outperform a simple random walk exchange rate model. These results are qualitatively

unaffected by the use of data revisions, by the use of a rolling—rather than recursive—forecasting scheme, and by the addition of an extra lag in the fundamentals.¹⁷

8 Discussion and Conclusions

This study focused on an empirical framework that explicitly takes into account the fact that there are changing weights in the economic fundamentals that, in theory, drive exchange rates, and that no model of fundamentals appears to perform well for long periods of time. Our framework allows for a menu of fundamentals that goes beyond the standard monetary fundamentals most commonly used in the literature since Mark (1995), comprising other variables suggested by exchange rate determination theory, including net foreign assets, interest rates differentials and the trade balance. Using a new set of real-time quarterly data spanning the period between 1977Q1 and 2003Q3, we employed a recursive procedure which selects, quarter by quarter, the best model on the basis of a variety of statistical and economic criteria within all possible combinations of fundamentals, allowing for the fundamentals model and the parameters to change over time.

Our main findings are as follows. First, the information embedded in the economic fundamentals, when used in a framework that allows for the set of fundamentals in the model to change over time and for parameter instability, can forecast the exchange rate with a remarkable degree of accuracy in some cases. However, this requires the ability to select the best model from the various set of models that can be used on the basis of information available, i.e. this requires that the investor must be have available a reliable model selection criterion to discriminate among different specifications of the fundamentals model. Second, if the conventional set of (statistical and economic) model selection criteria is used to pick up the best model from the large pool of models available, the same set of economic fundamentals is not useful in forecasting exchange rates out of sample. We have shown that models that optimally use the information in the fundamentals change often and this implies frequent shifts in the parameters. Standard model selection criteria appear to be unable to generate such shifts, yielding empirical exchange rate models that cannot forecast the exchange rate better than a random walk model. Therefore, while the stylized facts that fundamentals models cannot beat a no-change model is—yet again—confirmed in this paper, the reason for this result is, on the basis of the evidence presented here, different from many other studies. Our evidence suggests that the exchange rate disconnect puzzle arising from our result is unlikely to be caused by lack of information in the fundamentals, and more likely be the result of poor model selection criteria in this context.

¹⁷A further check we carried out involves estimating the final exchange rate equations in panel for all of the exchange rates examined in order to obtain more precise estimates of the parameters by exploiting the cross-correlation in the covariance matrix (see Mark and Sul, 2001). However, our results—not reported to conserve space—were qualitatively identical to the core results reported earlier.

These findings can be connected to much existing literature. Indeed, from a different angle, Engel and Hamilton (1990) recorded that long swings characterize the behavior of nominal exchange rates and, by taking them into account and estimating Markov-switching models, it is possible to generate better forecasts than a random walk. This result has been confirmed in other contributions where different fundamentals have been employed (see, *inter alia*, Clarida, Sarno, Taylor and Valente, 2003 and the references therein). However, these forecasting results appear to be somewhat fragile. Overall, in fact, the literature on nonlinear modelling of exchange rates has produced models that albeit fitting satisfactorily in sample, generally fail to beat simple random walk models or linear specifications in out-of-sample forecasting (e.g. see Diebold and Nason, 1990; Engel, 1994; Meese and Rose, 1990, 1991). This last evidence can be interpreted in light of our empirical results in that standard model selection criteria appear to be unable to detect and correctly identify shifts in the state variables (i.e. economic fundamentals). In fact, the existing empirical literature has focused primarily on the identification of long-swings in the exchange rate data. Our results suggest that in order to enhance the forecasting performance of predictive-regressions it is necessary to combine long-swings with short and recurrent swings in the predictive variables. Unfortunately, to date, econometric techniques able to thoroughly address this issue are not available and new techniques are awaited. Understanding the exact nature of these shifts and better ways to identify and model them remains an important challenge in this research agenda.

Table 1. Results from the ‘Virtual’ Out-of-Sample Exercise

		BM	RW
Japan	<i>MAE</i>	0.023 (0.008)	0.053 (–)
	<i>RMSE</i>	0.039 (0.019)	0.065 (–)
UK	<i>MAE</i>	0.017 (0.037)	0.039 (–)
	<i>RMSE</i>	0.032 (0.029)	0.051 (–)
Canada	<i>MAE</i>	0.012 (0.450)	0.019 (–)
	<i>RMSE</i>	0.020 (0.526)	0.026 (–)
Switzerland	<i>MAE</i>	0.030 (0.075)	0.054 (–)
	<i>RMSE</i>	0.048 (0.081)	0.068 (–)
Germany	<i>MAE</i>	0.026 (0.281)	0.052 (–)
	<i>RMSE</i>	0.040 (0.232)	0.064 (–)

Notes: MAE and RMSE denote mean absolute forecast errors and root mean square forecast errors respectively. BM is the best model chosen; RW is the random walk with drift model. Figures in parentheses are p -values of the null hypothesis of equal predictive accuracy between BM and RW. p -values are computed by using the White’s (2000) Reality Check procedure as indicated in the text and described in Appendix B.

Table 2. Independence of the Inclusion Frequencies

	ALL	Δs_{t-1}	mf_{t-1}	$(i - i^*)_{t-1}$	NFA_{t-1}	tb	tb^*	c
$m = 3$								
Japan	0	0	0.034	0.551	0	0.959	0.204	0.033
UK	0	0.333	0.059	0	0	0.267	0.001	0.774
Switzerland	0	0	0	0	0	0.896	–	0
$m = 4$								
Japan	0	0	0.030	0.946	0	0.660	0	0.003
UK	0	0.470	0	0.977	0	0	0	0.835
Switzerland	0	0.003	0	0	0.527	0.626	–	0

Notes: The figures reported are p -values of the BDSL test for the null hypothesis that the switches relative to the variables in each column are iid (Brock, Dechert, Scheinkman, LeBaron, 1996). The test statistics, and the relative p -values, are computed by setting the maximum value of embedding dimensions $m = 3, 4$ and the distance ϵ so as to ensure that 70% the total number of pairs of points in the sample lie within of each other (see Brock, Hsieh and Lebaron, 1991; Brock, Dechert, Scheinkman, LeBaron, 1996 for further details). ALL denotes the p -value for the null hypothesis that the switches are jointly independent; these p -values are calculated by using Bonferroni bounds as discussed in the text. 0 denotes a p -value over than 10^{-8} .

Table 3. Results from the Real-Time Out-of-Sample Exercise: Statistical criteria of evaluation

	MAE	RMSE	LAE	\bar{R}^2	AIC	BIC	SIGN	PIC	FIC	RW
Japan	<i>MAE</i>	0.052 (0.015)	0.057 (0.534)	0.052 (0.007)	0.051 (0.003)	0.056 (0.829)	0.056 (0.298)	0.050 (0.003)	0.050 (0.003)	0.053 (-)
	<i>RMSE</i>	0.068 (0.170)	0.071 (0.713)	0.067 (0.062)	0.066 (0.105)	0.071 (0.969)	0.068 (0.313)	0.065 (0.018)	0.065 (0.018)	0.065 (-)
UK	<i>MAE</i>	0.043 (0.678)	0.470 (0.995)	0.045 (0.939)	0.045 (0.973)	0.040 (0.754)	0.044 (0.908)	0.045 (0.900)	0.045 (0.900)	0.039 (-)
	<i>RMSE</i>	0.056 (0.682)	0.061 (0.996)	0.059 (0.945)	0.057 (0.917)	0.056 (0.949)	0.057 (0.903)	0.059 (0.920)	0.059 (0.920)	0.051 (-)
Switzerland	<i>MAE</i>	0.0546 (0.223)	0.056 (0.372)	0.054 (0.075)	0.058 (0.769)	0.054 (0.404)	0.054 (0.110)	0.053 (0.046)	0.053 (0.046)	0.054 (-)
	<i>RMSE</i>	0.072 (0.558)	0.073 (0.752)	0.068 (0.055)	0.071 (0.682)	0.069 (0.698)	0.068 (0.114)	0.068 (0.071)	0.068 (0.071)	0.068 (-)

Notes: MAE, RMSE and LAE denote mean absolute forecast error, root mean square error and the last absolute forecast error respectively. SIGN is a sign criterion based on the directional accuracy of the in sample predictions of the exchange rate change. AIC, BIC, PIC, FIC are the Akaike Information Criterion, the Schwarz Information Criterion, the Posterior Information Criterion and the Fisher Information Criterion respectively. RW denotes the random walk model of exchange rate. Columns denote real-time model selection criteria while rows indicate evaluation criteria. Figures in parentheses are *p*-values of the null hypothesis of equal predictive accuracy between the models selected according to the criterion indicated in each column and RW. *p*-values are computed by using the White's (2000) Reality Check procedure as indicated in the text and described in Appendix B.

Table 4. Results from the Real-Time Out-of-Sample Exercise: Economic criteria of evaluation

		MAE	RMSE	LAE	\bar{R}^2	AIC	BIC	SIGN	PIC	FIC	RW
Japan	EOPW	3.106 (0.486)	3.492 (0.599)	2.044 (0.218)	5.082 (0.904)	3.571 (0.505)	2.434 (0.737)	2.563 (0.309)	3.492 (0.588)	3.492 (0.588)	3.379 (-)
	RR	0.307 (0.512)	0.337 (0.592)	0.202 (0.267)	0.434 (0.814)	0.341 (0.596)	0.248 (0.241)	0.258 (0.369)	0.337 (0.581)	0.337 (0.581)	0.328 (-)
UK	EOPW	1.970 (0.257)	1.438 (0.147)	2.035 (0.251)	1.979 (0.242)	1.828 (0.163)	3.578 (0.677)	1.629 (0.924)	1.438 (0.149)	1.438 (0.149)	3.165 (-)
	RR	0.230 (0.255)	0.136 (0.119)	0.236 (0.260)	0.231 (0.263)	0.207 (0.150)	0.414 (0.649)	0.171 (0.123)	0.136 (0.125)	0.136 (0.125)	0.374 (-)
Switzerland	EOPW	2.075 (0.138)	3.694 (0.489)	1.361 (0.081)	3.800 (0.503)	1.780 (0.096)	2.038 (0.123)	4.038 (0.586)	3.694 (0.497)	3.694 (0.497)	3.964 (-)
	RR	0.201 (0.181)	0.338 (0.519)	0.105 (0.057)	0.346 (0.519)	0.167 (0.093)	0.197 (0.149)	0.358 (0.571)	0.338 (0.519)	0.338 (0.519)	0.357 (-)

Notes: EOPW and RR denote the recursive end-of-period wealth criterion and the recursive return criterion calculated as described in Section 3. See notes to Table 3.

Table 5. Results With Revised Data: Dollar-Yen

Panel A) 'Virtual' Out-of-Sample Exercise

	BM	RW
<i>MAE</i>	0.023	0.053
	(0.023)	(-)
<i>RMSE</i>	0.039	0.065
	(0.019)	(-)

Panel B) Real-Time Out-of-Sample Exercise

	MAE	RMSE	LAE	\bar{R}^2	AIC	BIC	SIGN	PIC	FIC	RW
<i>MAE</i>	0.053	0.051	0.056	0.052	0.052	0.056	0.054	0.052	0.052	0.053
	(0.013)	(0.007)	(0.591)	(0.009)	(0.004)	(0.732)	(0.304)	(0.004)	(0.002)	(-)
<i>RMSE</i>	0.069	0.066	0.071	0.067	0.067	0.070	0.065	0.066	0.066	0.065
	(0.152)	(0.021)	(0.672)	(0.064)	(0.112)	(0.884)	(0.356)	(0.015)	(0.016)	(-)

Notes: The results reported in this table replicate the ones in Tables 1 and 3 where historical (revised) data are used. See notes to Tables 1 and 3.

Table 6. Results Using Rolling Forecasting Scheme: Dollar-Yen

Panel A) 'Virtual' Out-of-Sample Exercise

	BM	RW
<i>MAE</i>	0.017	0.053
	(0.025)	(-)
<i>RMSE</i>	0.035	0.066
	(0.020)	(-)

Panel B) Real-Time Out-of-Sample Exercise

	MAE	RMSE	LAE	\bar{R}^2	AIC	BIC	SIGN	PIC	FIC	RW
<i>MAE</i>	0.067	0.068	0.065	0.064	0.064	0.061	0.063	0.068	0.068	0.053
	(0.014)	(0.006)	(0.673)	(0.008)	(0.008)	(0.389)	(0.354)	(0.003)	(0.003)	(-)
<i>RMSE</i>	0.085	0.085	0.082	0.080	0.079	0.077	0.078	0.085	0.085	0.066
	(0.195)	(0.024)	(0.605)	(0.067)	(0.132)	(0.791)	(0.383)	(0.012)	(0.015)	(-)

Notes: The results reported in this table replicate the ones in Tables 1 and 3 where a rolling forecasting scheme is used. The rolling window is set up to 5 years (i.e. 20 quarters) See notes to Tables 1 and 3.

Table 7. Results Using Extra Lags in the Fundamentals: Dollar-Yen

Panel A) 'Virtual' Out-of-Sample Exercise

	BM	RW
<i>MAE</i>	0.019	0.053
	(0.030)	(-)
<i>RMSE</i>	0.034	0.065
	(0.035)	(-)

Panel B) Real-Time Out-of-Sample Exercise

	MAE	RMSE	LAE	\bar{R}^2	AIC	BIC	SIGN	PIC	FIC	RW
<i>MAE</i>	0.050	0.048	0.054	0.049	0.051	0.053	0.054	0.048	0.048	0.053
	(0.015)	(0.020)	(0.496)	(0.005)	(0.007)	(0.404)	(0.630)	(0.004)	(0.008)	(-)
<i>RMSE</i>	0.066	0.064	0.069	0.064	0.067	0.070	0.068	0.064	0.064	0.065
	(0.205)	(0.027)	(0.550)	(0.078)	(0.164)	(0.702)	(0.421)	(0.011)	(0.013)	(-)

Notes: The results reported in this table replicate the ones in Tables 1 and 3 where estimations are carried out using an additional second lag in each of the regressors as discussed in Section 7. See notes to Tables 1 and 3.

A Appendix: Real Time Data Set

A.1 Macroeconomic Data

For the purposes of our analysis we assembled and compiled real time data of quarterly frequency from past issues of OECD's *Main Economic Indicators* (MEI). In particular, we constructed four vintages for each year considered starting from 1988 using the January, April, July and October issues of the MEI. This gives a total of 65 data vintages from January 1988 to January 2004 for Canada, United States (US), Japan, Germany, Switzerland and United Kingdom (UK). The January and October 1998 issues of the MEI were missing; therefore we used the February and September 1998 issues, respectively of the MEI. Each vintage consists of quarterly historic data of the variables considered until 1984:04 and the values reported at the corresponding MEI issue. The historic data covered the period 1973:01-1984:04 and were retrieved from the *Main Economic Indicators Historical Statistics 1969-1988* issue published by OECD.

For Canada, US, Japan, Germany, Switzerland and UK we used the Gross Domestic Product (GDP), expressed in the corresponding national currency, seasonally adjusted. We used the Consumer Price Index (CPI, all items) for Canada, Japan, Germany and Switzerland and the CPI (all items) for wage earners for the US.

We retrieved seasonally adjusted and unadjusted data for M1 were possible with respect to the countries considered expressed in the national currency unit. Data for the German M1 were available only until to the 1997:07 vintage, while only seasonally unadjusted data were available for Switzerland. For the UK we approximated the monetary aggregate by M1 plus quasi money up to the 1989:10 vintage and by M2 thereafter because of data availability.

Seasonally adjusted data for the current trade balance expressed in the national currency of the country examined were retrieved. Trade balance for Japan, Germany and Canada was seasonally adjusted by using the procedure followed by Mark and Sul (2001). Data for the trade balance with respect to Switzerland are first reported at the 1995:07 vintage with starting point 1993:01. Data for the Japanese trade balance were expressed in US dollars up to the 1996:07 vintage, although they were reported in Japanese yen at the subsequent vintages. Hence, we converted the corresponding data into Japanese yen using the US dollar - Japanese yen spot exchange rate reported in the MEI.

Data for Germany up to the 2002:01 vintage were expressed in German marks (DM) and to Euros for the subsequent vintages. Therefore, we used the 1999 conversion rate for the Deutsche mark (1 Euro = 1.95583 DM) in order to convert German data after the 2002:04 vintage in DM.

A.2 Data on International Capital Movements

With respect to international capital movements, we constructed data vintages for the US transactions with foreigners in long-term foreign and domestic securities using back issues of the *Treasury Bulletin* published quarterly in March, June, September and December each calendar year by the US Department of the Treasury. The data cover transactions carried out in the US for the accounts of foreigners and transactions executed abroad for the accounts of institutions and individuals resident in the US. These transactions involve Long-Term Domestic and Foreign Securities and they are classified by type and country of the foreign buyers and sellers who deal with the US reporting institutions. Each issue

reports the total purchases and sales by foreigners from and to US residents, respectively during the previous of the issue quarter.

Starting from the December 1987 to the December 2003 issue of the *Treasury Bulletin* we constructed four vintages per calendar year. This leaves us with 65 vintages. Each vintage for each country considered consists of quarterly historic data on total purchases and sales spanning the period 1977:01 to the quarter before the quarter of publication of the corresponding *Treasury Bulletin* issue and the values for total purchases and sales reported at that issue. Monthly historic data on capital movements are retrieved by the US Department of the Treasury’s web site.¹⁸ We converted these monthly historic data into quarterly by summing up the values in total purchases that correspond to the period between the issues of the *Treasury Bulletin*. The same task was performed for the total sales of foreign and domestic securities in the US in order to construct the necessary quarterly historic data for total sales.

B Appendix: Bootstrap procedure for the p -value of the test of equal forecast accuracy

We employ the Reality Check procedure proposed by White (2000), building on previous work by Diebold and Mariano (1995) and West (1996), for testing the null hypothesis that the best model selected in a specification search has no predictive superiority over a given benchmark model. This procedure is designed to allow aggressive model searching to be undertaken with confidence that one will not mistake results that could have been generated by chance for genuinely good results. The test evaluates the distribution of a performance measure accounting for the full set of models that lead to the best-performing model, and is based on the $L \times 1$ performance statistic

$$\bar{f} = N^{-1} \sum_{t=R}^T \hat{f}_{t+1}$$

where L is the number of “configurations” (or possible models, equal 128 per time period in this paper); N is the number of prediction periods indexed from R through T so that $T = R + N - 1$; \hat{f}_t is the observed performance measure for period t (such as the difference in RMSEs between the benchmark and a model), obtained on the basis of a model specification, set of predictor variables and parameter estimates. In our application, the various parameterizations of the models directly generate one-quarter-ahead forecasts of the exchange rate that are used to measure performance; $N = 43$ since the first set of models are estimated from 1977Q1 to 1987Q3, and the first forecast is for the change in the exchange rate from 1987Q3 to 1987Q4.

We are interested in testing the null hypothesis that the best model selected in a specification search has no predictive superiority over a given benchmark model—in our case the random walk model—which can be expressed as

$$H_0: \max_{i=1, \dots, L} \{E(f_i)\} \leq 0$$

¹⁸<http://www.ustreas.gov/tic>

Rejection of this null implies that the best model achieves performance that is superior to the random walk benchmark.¹⁹

White (2000) shows how this null hypothesis can be evaluated using the stationary bootstrap of Politis and Romano (1994) to the observed values of $f_{i,t}$. The bootstrap algorithm used to determine the p -values of the test statistic of equal point forecast accuracy consists of the following steps:

1. Given the sequence of observations $\{Y_{t+1}\}$ where $Y_{t+1} = (\Delta s_{t+1}, X_t)'$ and X_t denotes the explanatory variables, estimate each of the possible configurations under model (3) in Section 3 and construct the test statistic of interest, $\hat{\theta}$ (i.e. test statistic for the null hypothesis, H_0 , of equal point forecast accuracy relative to a random walk) for each configuration—the total number of configurations at each point in time is 128 (2^7).

2. Postulate a data generating process (DGP) for each of the configurations described in step 1, where the exchange rate is assumed to follow a driftless random walk under H_0 and the innovations are assumed to be iid

3. Based on the models specified in step 2), bootstrap with replacement to generate a sequence of pseudo observations $\{Y_{t+1}^*\}$ of the same length as the original data series $\{Y_{t+1}\}$. Repeat this step 1,000 times.

4. For each of the 1,000 bootstrap replications $\{Y_{t+1}^*\}$, estimate each of the configurations and construct the test statistic of interest, $\hat{\theta}^*$.

5. Construct the following statistics: $\bar{V} = \max_{i=1,\dots,L} \left\{ \sqrt{N} (\bar{f}_i) \right\}$ and $\bar{V}_i^* = \max_{j=1,\dots,L} \left\{ \sqrt{N} (\bar{f}_{j,i}^* - \bar{f}_i) \right\}$ where $j = 1, \dots, 1,000$ is the number of replications.

6. Use the empirical distribution of the 1,000 replications of the bootstrap test statistic, $\hat{\theta}^*$ to determine the p -value of the test statistic $\hat{\theta}$. In other words, the Reality Check p -value is obtained by comparing \bar{V} to the quantiles of \bar{V}_i^* . By employing the maximum value over all models, the p -value incorporates the effects of data snooping from the search over the L models at each point time.

¹⁹This null hypothesis is relevant, for example, for the case when the performance measure is the RMSE, where a lower number implies better performance. The inequality changes sign under the null hypothesis when the performance measure used improves for larger positive values, such as, for example, the RR (Sharpe ratio) criterion.

References

- Abhyankar, A., Sarno, L. and Valente, G. (2005), "Exchange Rates and Fundamentals: Evidence on the Economic Value of Predictability," *Journal of International Economics*, 66, 325-348.
- Amano, R.A. and van Norden, S. (1998), "Oil Prices and the Rise and Fall of the US Real Exchange Rate," *Journal of International Money and Finance*, 17, pp. 299-316.
- Andersen, T., Bollerslev, T., Diebold, F.X. and Vega, C. (2003), "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review*, 93, 38-62.
- Bacchetta, P. and Van Wincoop, E. (2004), "A Scapegoat Model of Exchange-Rate Fluctuations," *American Economic Review Papers and Proceedings*, 94, 114-118.
- Bacchetta, P. and Van Wincoop, E. (2005), "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?," *American Economic Review*, forthcoming.
- Backus, D., Gregory, A. and Telmer, C. (1993), "Accounting for Forward Rates in Markets for Foreign Currency," *Journal of Finance*, 48, 1887-1908.
- Baxter, M. and Stockman, A.C. (1989), "Business Cycles and the Exchange Rate Regime: Some International Evidence," *Journal of Monetary Economics*, 23, 377-400.
- Berkowitz, J. and Giorgianni, L. (2001), "Long-Horizon Exchange Rate Predictability?," *Review of Economics and Statistics*, 83, 81-91.
- Bossaerts, P. and Hillion, P. (1999), "Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?," *Review of Financial Studies*, 12, 405-428.
- Branson, W.H., (1984), "A Model of Exchange Rate Determination With Policy Reaction: Evidence from Monthly Data," in P. Malgrange and P.A. Muet (eds.) *Contemporary Macroeconomic Modelling*, Oxford: Basil Blackwell.
- Brock, W.A., Dechert, W.D., Scheinkman, J.A. and LeBaron, B. (1996), "A Test for Independence Based on the Correlation Dimension," *Econometric Reviews*, 15, 197-235.
- Brock, W.A., Hsieh, D. and LeBaron, B. (1991), *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*. MIT Press Cambridge, Massachusetts.
- Cavallo, M. and Ghironi, F. (2002), "Net Foreign Assets and the Exchange Rate: Redux Revived," *Journal of Monetary Economics*, 49, 1057-97.
- Chen, Y.-C. and Rogoff, K. (2003), "Commodity Currencies," *Journal of International Economics*, 60, pp. 133-160.
- Cheung, Y., and Chinn, M. 2001, "Currency Traders and Exchange Rate Dynamics: A Survey of the U.S. Market," *Journal of International Money and Finance*, 20, 439-471.
- Cheung, Y.-W., Chinn, M.D. and Pascual, A.G. (2003), "Recent Exchange Rate Models: In-Sample Fit and Out-of-Sample Performance," in: De Grauwe, P. (ed.), *Exchange Rate Modeling: Where Do Stand?*, Cambridge and New York: Cambridge University Press.
- Clarida, R.H., Sarno, L., Taylor, M.P. and Valente, G. (2003), "The Out-of-Sample Success of Term Structure Models as Exchange Rate Predictors: One Step Beyond," *Journal of International Economics*, 60, 61-83.

- Diebold, F.X. and Mariano, R.S. (1995), "Comparing Predictive Accuracy," *Journal of Business and Economics Statistics*, 13, 253-263.
- Diebold, F.X. and Nason, J. (1990), "Nonparametric Exchange Rate Prediction?," *Journal of International Economics*, 28, 315-322.
- Engel, C. (1994), "Can the Markov Switching Model Forecast Exchange Rates?," *Journal of International Economics*, 36, 151-165.
- Engel, C. (1996), "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence," *Journal of Empirical Finance*, 3, 123-92.
- Engel, C. and Hamilton, J.D., (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?," *American Economic Review*, 80, 689-713.
- Engel, C. and West, K.D. (2005), "Exchange Rates and Fundamentals," *Journal of Political Economy*, 113, 485-517.
- Faust, J., Rogers, J.H. and Wright, J. (2003), "Exchange Rate Forecasting: The Errors We've Really Made," *Journal of International Economics*, 60, 35-59.
- Flood, R.P. and Rose, A.K. (1995), "Fixing Exchange Rates A Virtual Quest for Fundamentals," *Journal of Monetary Economics*, 36, 3-37.
- Flood, R.P. and Rose, A.K. (1996), "Fixes: Of the Forward Discount Puzzle," *Review of Economics and Statistics*, 78, 748-52.
- Flood, R.P. and Rose, A.K. (1999), "Understanding Exchange Rate Volatility Without the Contrivance of Macroeconomics," *Economic Journal*, 109, 660-72.
- Frankel, J.A. (1996), "How Well Do Foreign Exchange Markets Function: Might a Tobin Tax Help?," in ul Haq, M., Kaul, I. and Grunberg, I. (eds.), *The Tobin Tax: Coping with Financial Volatility*, Oxford University Press, New York, 41-81.
- Frankel, J. and Froot, K.A. (1987), "Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations," *American Economic Review*, 77, 133-53.
- Frenkel, J.A. (1976), "A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence," *Scandinavian Journal of Economics*, 78, 200-24.
- Frenkel, J.A. and Johnson, H.A. (1978), *The Economics of Exchange Rates: Selected Studies*, Reading, Mass.: Addison-Wesley.
- Galambos, J. and Simonelli, I. (1996), *Bonferroni - type Inequalities with Applications*, Springer, New York.
- Groen, J.J.J. (2000), "The Monetary Exchange Rate Model as a Long-Run Phenomenon," *Journal of International Economics*, 52, 299-319.
- Hunter, D. (1976), "An Upper Bound for the Probability of a Union," *Journal of Applied Probability*, 13, 597-603.
- Kilian, L. (1999), "Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?," *Journal of Applied Econometrics*, 14, 491-510.
- Kilian, L. and Taylor, M.P. (2003), "Why is it So Difficult to Beat the Random Walk Forecast of Exchange Rates?," *Journal of International Economics*, 60, 85-107.

- Kounias, S. (1968), "Bounds for the Probability of the Union With Applications," *Annals of Mathematical Statistics*, 39, 2154-2158.
- Kreuger, A.D. (1983), *Exchange Rate Determination*, Cambridge University Press.
- Kumhof, M. and Van Nieuwerburgh, S. (2005), "Monetary Policy in an Equilibrium Portfolio Balance Model," International Monetary Fund and New York University, mimeo.
- Lane, P.R. (2001), "The New Open Economy Macroeconomics: A Survey," *Journal of International Economics*, 54, 235-66.
- Levich, R.M. and Thomas, L. (1993), "The Significance of Technical Trading-Rule Profits in the Foreign Exchange Market: A Bootstrap Approach," *Journal of International Money and Finance*, 12, 451-474.
- Lucas, R.E. Jr. (1976), "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, 19-46.
- Lucas, R.E. Jr. (1982), "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics*, 10, 335-59.
- Mark, N.C. (1995), "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85, 201-18.
- Mark, N.C. and Sul, D. (2001), "Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel," *Journal of International Economics*, 53, 29-52.
- Meese, R.A. and Rogoff, K. (1983), "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics*, 14, 3-24.
- Meese, R.A. and Rose, A.K. (1990), "Nonlinear, Nonparametric, Nonessential Exchange Rate Estimation," *American Economic Review*, 80, 192-96.
- Meese, R. and Rose, A. (1991), "An empirical Assessment of Non-Linearities in Models of Exchange Rate Determination," *The Review of Economic Studies*, 58, 603-619.
- Mussa, M. (1976), "The Exchange Rate, the Balance of Payments and Monetary and Fiscal Policy under a Regime of Controlled Floating," *Scandinavian Journal of Economics*, 78, 229-48.
- Mussa, M. (1979), "Empirical Regularities in the Behavior of Exchange Rates and Theories of the Foreign Exchange Market," *Carnegie-Rochester Conference Series on Public Policy*, 11, 9-57.
- Mussa, M. (1986), "Nominal Exchange Rate Regimes and the Behavior of Real Exchange Rates: Evidence and Implications," in Karl Brunner and Allan H. Meltzer, eds., *Carnegie-Rochester, Conference Series on Public Policy: Real Business Cycles, Real Exchange Rates, and Actual Policies*, 25, 117-214.
- Neely, C.J. and Sarno, L. (2002), "How Well Do Monetary Fundamentals Forecast Exchange Rates?," *Federal Reserve Bank of St. Louis Review*, 84, 51-74.
- Neely, C.J., Weller, P. and Dittmar, R. (1997), "Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach," *Journal of Financial and Quantitative Analysis*, 32, 405-26.
- Ng, S. and Perron, P. (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 69, 1519-1554.
- Obstfeld, M. (2004), "External Adjustment," *Review of World Economics*, forthcoming.

- Obstfeld, M. and Rogoff, K. (1995), "Exchange Rate Dynamics Redux," *Journal of Political Economy*, 103, 624-60.
- Obstfeld, M. and Rogoff, K. (2000), "The Six Major Puzzles in International Macroeconomics: Is there a Common Cause?," in: Bernanke, B. and Rogoff, K. (eds.), *NBER Macroeconomics Annual 2000*, Cambridge, Mass: MIT Press, 339-90.
- Obstfeld, M. and Rogoff, K. (2004), "The Unsustainable US Current Account Position Revisited," *National Bureau of Economic Research Working Paper*, 10869.
- Pesaran, H.M. and Timmermann, A. (1995), "Predictability of Stock returns: Robustness and Economic Significance," *Journal of Finance*, 50, 1201-1228.
- Pesaran, H.M. and Timmermann, A. (2000), "A Recursive Modelling Approach to Predicting UK Stock Returns," *Economic Journal*, 110, 159-191.
- Phillips, P.C.B. and Ploberger, W. (1996), "Posterior Odds Testing for a Unit Root With Data Based Model Selection," *Econometrica*, 64, 381-412.
- Politis, D.N. and Romano, J.P. (1994), "The Stationary Bootstrap," *Journal of the American Statistical Association*, 89, 1303-1313.
- Rapach, D.E. and Wohar, M.E., (2002), "Testing The Monetary Model of Exchange Rate Determination: New Evidence From a Century of Data," *Journal of International Economics*, 58, 359-85.
- Rose, A.K. and Yellen, J.L. (1989), "Is there a J-curve?," *Journal of Monetary Economics*, 24, 53-68.
- Rossi, B. (2005a), "Are Exchange Rates Really Random Walks? Some Evidence Robust to Parameter Instability," *Macroeconomic Dynamics*, forthcoming.
- Rossi, B. (2005b), "Testing Long-Horizon Predictive Ability With High Persistence, and the Meese-Rogoff Puzzle," *International Economic Review*, forthcoming.
- Sarno, L. and Taylor, M.P. (2001), *The Microstructure of the Foreign Exchange Market*, Princeton Studies in International Economics, 89, International Economics Section, Princeton, NJ: Princeton University Press.
- Schinasi, G.J. and Swamy, P.A.V.B. (1989), "The Out-of-Sample Forecasting Performance of Exchange Rate Models When Coefficients Are Allowed to Change," *Journal of International Money and Finance*, 8, 375-90.
- Sullivan, R., Timmerman, A. and White, H. (1999), "Data-Snooping, Technical Trading Rule Performance, and the Bootstrap," *Journal of Finance*, 54, 1647-1692.
- Wei, C. (1992), "On Predictive Least Squares Principles," *Annals of Statistics*, 20, 1-42.
- West, K.D. (1996), "Asymptotic Inference about Predictive Ability," *Econometrica*, 64, 1067-1084.
- West, K.D. and McCracken, M.W. (1998), "Regression-Based Tests of Predictive Ability," *International Economic Review*, 39, 817-40.
- White, H. (2000), "A Reality Check For Data Snooping," *Econometrica*, 68, 1097-1127.