Heterogeneous Agents, the Financial Crisis and Exchange Rate Predictability*

Daniel Buncic
Institute of Mathematics & Statistics
University of St. Gallen
Switzerland

Gion Donat Piras
School of Economics & Political Science
University of St. Gallen
Switzerland

First Version: October 2, 2014
This Version (fx.3e): March 23, 2015

Abstract

We construct an empirical heterogeneous agent model which optimally combines forecasts from fundamentalist and chartists agents and evaluate its out-of-sample forecast performance using daily date covering the period from January 1999 to June 2014 for six of the most widely traded currencies. We use daily financial data such as level, slope and curvature yield curve factors, equity prices, as well as risk aversion and global trade activity measures in the fundamentalist agent's predictor set to obtain a proxy the market's view on the state of the macroeconomy. Chartist agents rely upon standard momentum, moving average and relative strength index indicators in their predictor set. The individual agent specific forecasts are computed using the recently proposed flexible dynamic model averaging framework and are then aggregated into a model combined forecast using a forecast combination regression. We show that our empirical heterogeneous agent model produces statistically significant and sizable forecast improvements over the standard random walk benchmark, reaching out-of-sample R^2 values of 1.41, 1.07, 0.99 and 0.74 percent at the daily one-step ahead horizon for 4 out of the 6 currencies that we consider. Forecast gains remain significant for horizons up to three-days ahead. We show further that for 5 out of the 6 currencies, a substantial part of the forecast gains are realised over the September 2008 to February 2009 period, that is, around the time of the Lehman Brothers collapse. The time series evolution of the dynamic model combination weights shows that for the first half of the out-of-sample evaluation period, fundamentalist agents dominated the combination forecasts, while the last third of the out-of-sample period was driven by chartist agents.

Keywords: Empirical heterogeneous agent model, forecasting, time varying parameter model, state-space modelling, model combination, exchange rate predictability, financial crisis.
JEL Classification: C22, C52, C53, E17, F31, G17.

*We are grateful to Andrea Vedolin (LSE) for providing the yield curve factor loadings for Japan and Switzerland. We would like to thank Adrian Pagan, Paul Söderlind, Michael Bauer, Valentyn Panchenko and Andreas Schrimpf for helpful comments and discussions on earlier results in the paper.

Corresponding Author: Institute of Mathematics and Statistics, Bodanstrasse 6, 9000 St. Gallen, Switzerland. Tel: +41 (71) 224 2604. Email: daniel.buncic@unisg.ch. Web: http://www.danielbuncic.com.
Email: donatpiras@gmail.com.

1. Introduction

Since the seminal work of Meese and Rogoff (1983), it is well known that standard macroeconomic models of exchange rate determination have difficulties in producing *'significantly'* better forecasts than a simple random walk model. This finding is known as the *'Meese and Rogoff Puzzle'* in the exchange rate literature.[1]

Although there exists considerable disagreement about exchange rate predictability in the literature (see the recent and extensive review in Rossi (2013) for a discussion), empirical evidence seems to suggest that traditional economic predictors, such as interest rate, price and monetary differentials have rather weak predictive ability, especially when used in linear forecasting models and when considering short forecast horizons.[2] Rossi (2013) documents that the predictive ability of fundamental variables varies noticeably across currency pairs, the type of models that are used, and over the various sample periods that are considered in the literature. So far, there exists no evidence to indicate that any one set of predictors (or models) systematically outperforms the random walk model across various exchange rates and/or sample periods.

In this study, we take an entirely different modelling approach to the mainstream exchange rate forecasting literature. First, we construct an empirical heterogeneous agent model consisting of fundamentalist and chartist agents to form a model combined exchange rate forecast. More specifically, we use a simple model averaging approach to optimally weight the forecasts from the two individual agent types using a Granger and Ramanathan (1984) *'combination regression'*. Second, we use daily data to compute the agent specific forecasts. For chartist forecasts, this is a natural sampling frequency to adopt, as the most commonly used technical indicators such as moving average, momentum and relative strength index rules utilise daily data. For fundamentalist agents, which construct their forecasts based on macroeconomic information, we use daily financial variables as *'proxy variables'* to obtain information about the

[1] The term *'significant'* here is put in italics to emphasise that it is meant to be interpreted in the most general way, including economic and/or statistical significance.

[2] Some exceptions to this seem to be Monetary model based predictors when used to forecast multiple periods ahead and the recently applied variants of Taylor-rule based models and net foreign assets predictors.

state of the macroeconomy as it is perceive by financial market participants. The financial proxy variables that we use are yield curve data, stock price data, and data related to risk aversion and global trade activity. Third, we construct *individual agent specific* forecasts using the recently proposed Dynamic Model Averaging (henceforth DMA) framework. The DMA framework is an extremely flexible modelling approach that combines time varying parameters and model averaging into one methodological framework. Forecasting agents are known to switch and adapt their prediction models over time. In order to capture this stylised behaviour of agents, it is necessary to employ an econometric model that is able to mimic this flexibility.

So far, nearly all studies on exchange rate forecasting have relied upon low frequency data, ie., monthly or quarterly intervals, in order to be able to use standard macroeconomic fundamentals such as output, inflation rates, interest rates, etc., as predictors.[3] In this study, we want to take a different approach. Our objective is to construct forecasts that are as close to a '*real time*' forecasting scenario as possible. To achieve this, we use observed daily financial data to have a proxy for the '*market's view*' on current and future macroeconomic conditions. In this respect, we are mimicking a '*real time*' forecasting scenario, where only data that is observed up to time period t is used to construct a forecast $h-$steps ahead. The advantage of using daily financial data as a proxy for macroeconomic fundamentals is that we are able to avoid any ambiguities with respect to the timing of data releases (or its real time availability) and the impact of data revisions on the results of our study. Two well known weaknesses with using standard macroeconomic data for real time forecasting of exchange rates are: i) substantial data revisions and ii) considerable release lags in the data. These make a '*real time*' implementation and evaluation of exchange rate forecasts infeasible. The use of daily financial data has the benefit of providing a clear time stamp as to what data were actually available to the forecasting agent at the time the forecast was formed, without the potential drawback of having been revised.

Our empirical heterogeneous agent model allows us to include various technical indicators in the chartist predictor set in addition to the macroeconomic fundamentals. The use of

[3]See, for instance, Meese and Rogoff (1983), Mark (1995), Cheung *et al.* (2005), Wright (2008) for a few classic studies, and Rossi (2013) for a recent major review of the literature.

technical indicators as predictor variables has received considerably less attention in the recent empirical exchange rate modelling literature, despite their well documented and widespread use among practitioners (see, for instance, Allen and Taylor (1990), Taylor and Allen (1992), Menkhoff (1998), and Lui and Mole (1999) among many others). '*Chartists*', that is, agents that use technical indicators as trading signals, play also a crucial role in the theoretical finance literature on non-linear (dynamic) heterogeneous agent models (see, for instance, the classic papers by Brock and Hommes (1997, 1998) and also the ones by De Grauwe and Grimaldi (2005, 2006)). Neely *et al.* (2014) have recently shown that technical indicators can be used to improve forecasts of the equity risk premium. Moreover, using a model combination approach to average the individual agent specific predictions has two advantages over simply combining the fundamentalist and chartist regressors into one large joint predictor set. First, it substantially reduces the computational burden of the construction of the DMA forecasts, which requires the computation of all possible (linear) model combinations at each point in time. Second, it allows us to provided insights about the time varying '*importance*' of each agent's forecast in the model combined predictions.

One of the biggest difficulties when formulating a traditional fundamentalist forecasting model for exchange rates is that of model and parameter variability.[4] That is, not only are the parameters that link the expected change in the exchange rate to its fundamental value unstable over time, but more importantly, also the macroeconomic model that determines the fundamental value (or its predictors) is unlikely to remain the same over the span of the sample. For example, interest rate differentials, as specified by an Uncovered Interest Parity (UIP) condition, may be important predictors of exchange rates for a given period of time, but price or inflation differentials, making up the Purchasing Power Parity (PPP) relation, may be important for another sub-period of the sample.[5] In addition to the uncertainty as to what fundamental

[4]Note here that the same argument can be made for chartist agents. Chartists will switch their trading strategies between various different technical trading rules over time, depending on the profitability of the rule. This same argument is embedded in the theoretical heterogeneous agent models of Brock and Hommes (1997, 1998) and De Grauwe and Grimaldi (2005, 2006). For instance, momentum may be used for one period of time, while there could be a switch to using a resistance type indicator following a reversal trading strategy, once it is realised that momentum does not generate profits anymore.

[5]See also the discussion in Section 3 in Rossi (2013) for many more '*fundamental*' models and corresponding

model and/or variables to use in the specification, it is further well known that exchange rates are episodically unstable, may show *'long swings in the data'* and be *'disconnected'* from fundamentals, raising the need for a non-linear model specification.[6] Nevertheless, despite the need for a non-linear model, empirical out-of-sample forecast evaluations in favour of non-linear exchange rate models are rather weak.[7] Part of the empirical failure of non-linear exchange rate models is due to over-parameterisation, not allowing for breaks in the model structure and the use of only one *'fundamental'* predictor set in the construction of the forecasts.

In the existing literature, two main approaches are used to approximate unknown forms of non-linearity and model uncertainty. The first approach allows the parameters of a linear model specification to evolve over time. For instance, in the Time Varying Parameter (TVP) model of Stock and Watson (1998), the evolution of the parameters is assumed to follow a random walk process. The second approach is to approximate model uncertainty by averaging over or combining a large number of linear models. This is the philosophy underlying the Bayesian Model Averaging (BMA) literature. Forecasts from all possible (linear) model combinations are *'aggregated'* using the estimated posterior prediction probabilities. BMA was recently used successfully by Wright (2008) to predict exchange rates with a large set of potential predictors. Other studies that have used BMA in the finance literature, are Aramov (2002), Cremers (2002), Koop and Potter (2004) and Buncic and Melecky (2014). To allow for the most flexible econometric specification of the individual agent specific forecasts, we employ the DMA framework which combines time varying parameters and model averaging into one unified framework.

In this study, we use daily data for six of the most frequently traded currencies (relative to the US Dollar) to assess the out-of-sample forecast performance of our proposed empirical heterogeneous agent model. The six currencies are the Euro, the Yen, the British Pound, the Australian and Canadian Dollars, and the Swiss Franc. Using an out-of-sample evalua-

predictors.

[6]See Engel and Hamilton (1990), Rogoff (1996), and more recently, Altavilla and Grauwe (2010), who find that the switching nature of the exchange rate process is inconsistent with a linear representation.

[7]See, for instance, Klaassen (2005), who evaluates the Markov-Switching model of Engel and Hamilton (1990) and Buncic (2012) who studies the forecast performance of smooth transition type regime models of real exchange rates.

tion period from September 2001 to June 2014, we show that forecasts from the heterogeneous agent model significantly outperform the forecasts from the standard random walk benchmark model for all 6 currencies that are considered. More specifically, the Campbell and Thompson (2008) out-of-sample R^2 values corresponding to daily one-step ahead forecasts can be as high as 1.41%, 1.07%, 0.99%, and 0.74% for the Franc, the Euro, the Pound and the Yen series, and are somewhat lower for the Australian and Canadian Dollars at 0.29% and 0.24%. Additionally, standard statistical tests show that these forecast improvements are significant at the 10% level for the Australian and Canadian Dollars, and at the 1% level for the Euro, Yen, Pound and Franc. Forecast gains relative to the random walk benchmark remain statistically significant at the 10% level for forecasts up to three days ahead for the CHF, CAD, AUD, and EUR series, producing (three day) out-of-sample R^2 values as high as 0.33% and 0.30% for the Euro and Canadian Dollar, respectively.

To gain a better understanding of the statistical forecast evaluation results that we obtain, we follow the literature on forecasting the equity premium and investigate the evolution of the cumulative difference between the mean squared forecast errors of the random walk model and the model combined heterogeneous agent predictions over time. The time series plots of this cumulative difference shows a number of interesting and previously undocumented features in the predictive performance of the model. First, the effect of the Lehman Brothers collapse around the September 2008 to February 2009 period had a profound and largely positive effect on the predictive performance of the model combined heterogeneous agent forecasts. This result is visible for all currencies from the substantial upward movement in the cumulative difference of the mean squared forecast errors around the time of the Lehman Brothers collapse and is contrary to the findings in Adrian *et al.* (2010), Molodtsova *et al.* (2011), and Molodtsova and Papell (2012), who report a break down in predictive performance after this period. Second, this cumulative difference is positive for all six exchange rates over the entire out-of-sample period, with the only exception being the AUD series, for which from June 2010 onwards this sequence becomes negative. Overall, the effect of the Lehman Brothers collapse remains influential for forecast horizons up to 5 days ahead, being particularly visible for the Euro, the British Pound and the Canadian Dollar.

In order to learn about the time varying influence of chartist agents on the predictive performance of the heterogeneous agent model, we also investigate the time series evolution of the model combination weights. The combination weights show that for large parts of the first half of the out-of-sample evaluation period, mainly fundamentalist agents or macroeconomic predictor variables were contributing to the improvement in the model combined forecasts over the random walk model. Nevertheless, for the last third of the out-of-sample period, the influence of chartist forecasts increased noticeably, particularly for the GBP, EUR and AUD exchange rate series, and to a lesser extent also for the CHF and JPY. Over this period, these currencies experienced a fairly strong and rapid appreciation relative to the US Dollar, generating a considerable amount of momentum in currency returns.

Our contributions to the exchange rate literature can be summarised as follows. First, we construct an empirical heterogeneous agent model consisting of fundamentalist and chartist agents to forecast six of the most frequently traded currencies. Second, we use daily financial data as proxy variables for macroeconomic fundamentals to be able to construct daily real time fundamentalist forecasts. Third, we show that there exists a statistically significant and sizeable predictive component in daily exchange rate returns that is valid for forecast horizons up to 3 days ahead. Fourth, we highlight visually the important positive influence the period around and after the Lehman Brothers collapse has on the performance of the agent based model combined forecasts. And fifth, we show how the weight of chartists in the model combined heterogeneous agent predictions varies over time, with the weight function on chartists increasing in the last third of the out-of-sample period.

The remainder of the paper is organized as follows. Section 2 describes the model combination approach that we adopt in the paper. Section 3 describes in detail the fundamental and technical indicator data used by the two different types of agents of interest. Section 4 presents the empirical out-of-sample forecast evaluation results, together with a discussion. Section 5 concludes the paper with a summary and potential future research topics.

2. Modelling approach

We use a model combination (or averaging) approach to forecast the evolution of the exchange rate. This is implemented by combining the out-of-sample forecasts from two different types of agents which use two different sets of predictor variables to form their forecasts. The two agent types are *i*) *fundamentalists* and *ii*) *chartists*. Fundamentalists use variables that provide information about the *'strength'* of the underlying macroeconomy to construct forecasts of the exchange rate. Chartists, on the other hand, rely upon technical indicators (or trading rules) which give an indication of momentum and trend following behaviour in exchange rates, as well as oversold or overbought conditions.[8] Our intention here is to mimic the empirical behaviour of heterogeneous agents in foreign exchange markets, by first constructing forecasts of each individual agent type and then aggregating these two forecasts (optimally) by means of a Granger and Ramanathan (1984) forecast combination regression. It is well known since the seminal work of Bates and Granger (1969) that more accurate forecasts can be obtained by combining forecasts from several models.[9]

The model combined forecast from the two agents' individual forecasts are constructed as follows. Denote by y_{t+1} the variable to be predicted at time $t+1$ (ie., for simplicity of exposition, we can just think of a one-step ahead forecast here, but this is easily generalised to any forecast horizon h), and let $\hat{y}^{\mathsf{F}}_{t+1|t}$ and $\hat{y}^{\mathsf{C}}_{t+1|t}$ be the forecasts constructed by our two agents, fundamentalists and chartists, respectively. The model combined (MC) forecast is then obtained as a weighted average of the fundamentalist and chartist predictions, that is:

$$\hat{y}^{\mathsf{MC}}_{t+1|t} = (1-\hat{\omega}_t)\hat{y}^{\mathsf{F}}_{t+1|t} + \hat{\omega}_{t,h}\hat{y}^{\mathsf{C}}_{t+1|t} \tag{1}$$

where $\hat{\omega}_t$ is the fitted value from a constrained least squares regression of the form:

$$y_t = (1-\omega_t)\hat{y}^{\mathsf{F}}_{t|t-1} + \omega_t\hat{y}^{\mathsf{C}}_{t|t-1} + \nu_t, \tag{2}$$

[8] Exact details how these are constructed are given in Section 3.2.

[9] See also Hansen (2008) for a recent review of model averaging estimators.

with $\hat{y}^{\mathsf{F}}_{t|t-1}$ and $\hat{y}^{\mathsf{C}}_{t|t-1}$ being respectively the time t forecasts of y_t using information at time $t-1$, and ν_t is an error term. Note here, that since we are using the time t estimate of ω_t to forecast y_{t+1}, this effectively implies that we assume a random walk evolution for ω_t. Also, in the empirical implementation of the forecast combination, we re-estimate ω_t for each new observation that becomes available using a rolling window scheme. This effectively implies that a time-varying $\hat{\omega}_t$ is used in the forecast combination. Using a rolling window does not only give a more accurate representation of the real time forecasting behaviour of foreign exchange agents, but has the added benefit of providing the time-varying weights of the optimally combined forecasts and therefore the *'influence'* or *'activity'* of the two agent types that we model.

2.1. Constructing Fundamentalist and Chartist forecasts

In order to construct the model combined forecast from the two different agent types in (1), *individual* fundamentalist and chartist forecasts, that is, $\hat{y}^{\mathsf{F}}_{t+1|t}$ and $\hat{y}^{\mathsf{C}}_{t+1|t}$ are needed. A key feature of the agent specific forecasts is that they will be constructed from flexible models that evolve over time, because forecasting agents tend to re-estimate (or re-calibrate) their prediction models as new information becomes available. More importantly, it seems also highly likely that the set of predictors (or models) used in the construction of the forecasts will change over time. This could be due to the individuals which construct the forecast changing over time as a consequence of staff turnover.[10] Alternatively, an agent may prefer to construct forecasts from various *simple* models which could then be averaged, with the averaging weights being based on some preferred prediction optimality criterion specified by the agent. To be able to mimic the individual forecasts from our two agents as accurately and as flexibly as possible, we use the recently proposed Dynamic Model Averaging (DMA) framework.[11] What makes the DMA framework particularly appealing in the given context is its combination of time varying parameters and model averaging into one unifying framework, therefore mimicking

[10] For example, one chartist agent working for a firm in one time period may construct a trading strategy based on momentum or trend following, while another the comes to fill her position could prefer trading on reversals, ie., on overbought or oversold signals.

[11] The DMA framework was first introduced by Raftery *et al.* (2010) and applied to inflation forecasting in an economic context by Koop and Korobilis (2012).

the behaviour of agent learning and updating over time.

To outline how the DMA framework is implemented, let y_t denote the variable to be predicted at time period t.[12] Also, let $\mathbf{x}_{t-1}$ be a $(1 \times K)$ vector that contains the full set of k predictors plus an intercept term ($K = k + 1$), and let $m = 1, \ldots, M$ denote the model index, where $M = 2^k$ is the total number of possible (linear) model combinations (including the trivial model with only a constant term in it).[13] The set of predictors contained in the m^{th} model is denoted by $\mathbf{x}_{t-1}^{(m)}$, with the dimension of $\mathbf{x}_{t-1}^{(m)}$ being $(1 \times K_m)$. The two equations that make up the DMA framework (for model m) are:

$$\text{Measurement}: \underset{(1\times1)}{y_t} = \underset{(1\times K_m)}{\mathbf{x}_{t-1}^{(m)}} \underset{(K_m\times1)}{\boldsymbol{\beta}_t^{(m)}} + \underset{(1\times1)}{u_t^{(m)}} \tag{3a}$$

$$\text{State}: \underset{(K_m\times1)}{\boldsymbol{\beta}_t^{(m)}} = \underset{(K_m\times1)}{\boldsymbol{\beta}_{t-1}^{(m)}} + \underset{(K_m\times1)}{\boldsymbol{\epsilon}_t^{(m)}}, \tag{3b}$$

where (3a) and (3b) are measurement and state equations, respectively. The two disturbance terms $u_t^{(m)}$ and $\boldsymbol{\epsilon}_t^{(m)}$ in (3) are jointly Multivariate Normal (MN) distributed, uncorrelated with each other and over time, that is:

$$\begin{bmatrix} u_t^{(m)} \\ \boldsymbol{\epsilon}_t^{(m)} \end{bmatrix} \sim \text{MN}\left(\begin{bmatrix} 0 \\ \underset{(K_m\times1)}{\mathbf{0}} \end{bmatrix}, \begin{bmatrix} \underset{(1\times1)}{H_t^{(m)}} & 0 \\ \underset{(K_m\times K_m)}{\mathbf{0}} & \underset{(K_m\times K_m)}{\boldsymbol{Q}_t^{(m)}} \end{bmatrix} \right), \tag{4}$$

where $H_t^{(m)}$ and $\boldsymbol{Q}_t^{(m)}$ are the variance and covariance matrix of the measurement and state equations, respectively.

Also, let $\mathcal{M}_t$ denote the set of all possible models at time t, so that $\mathcal{M}_t \in \{1, 2, \ldots, M\}$. Given knowledge of $H_t^{(m)}$ and $\boldsymbol{Q}_t^{(m)}$ and by fixing the model set $\mathcal{M}_t = m$, ie., to one particular model,

[12] For reasons of simplicity, we use standard y_t and $\mathbf{x}_t$ notation to denote the left-hand side and predictor variables in the general description of the modelling framework. In our setting, y_t is the daily exchange rate return. This will be made explicit in Section 3 and Section 4, where the data and the forecast evaluation results are discussed.

[13] The term model here refers to the different possible linear combinations that can be obtained from using k predictors in a regression context, rather than the more general definition, where a model can be anything, potentially as flexible as non-linear or a non-parametric specification. The use of the term model is standard in the model averaging literature.

the system in (3) takes the form of a standard state-space model, making it thereby possible to extract or *'filter'* the time varying parameters $\boldsymbol{\beta}_t^{(m)}$ as the *'latent states'* using standard Kalman Filter recursions. One-step ahead forecasts and forecast errors are available as a by product of the Kalman Filter. Given $\mathcal{M}_t = m$, $H_t^{(m)}$ and $\boldsymbol{Q}_t^{(m)}$, the Kalman Filter recursions are:

$$\begin{aligned}
\text{Prediction}: \quad \hat{\boldsymbol{\beta}}_{t|t-1}^{(m)} &= \hat{\boldsymbol{\beta}}_{t-1|t-1}^{(m)} \\
\boldsymbol{P}_{t|t-1}^{(m)} &= \boldsymbol{P}_{t-1|t-1}^{(m)} + \boldsymbol{Q}_t^{(m)} && \text{(5a)} \\
\hat{y}_{t|t-1}^{(m)} &= \mathbf{x}_{t-1}^{(m)} \hat{\boldsymbol{\beta}}_{t|t-1}^{(m)} && \text{(5b)} \\
\text{Prediction errors}: \quad \hat{u}_t^{(m)} &= (y_t - \hat{y}_{t|t-1}^{(m)}) \\
\text{MSE of prediction errors}: \quad F_t^{(m)} &= \mathbf{x}_{t-1}^{(m)} \boldsymbol{P}_{t|t-1}^{(m)} \mathbf{x}_{t-1}^{\top(m)} + H_t^{(m)} && \text{(5c)} \\
\text{Kalman Gain}: \quad \boldsymbol{G}_t^{(m)} &= \boldsymbol{P}_{t|t-1}^{(m)} \mathbf{x}_{t-1}^{\top(m)} / F_t^{(m)} \\
\text{Updating}: \quad \hat{\boldsymbol{\beta}}_{t|t}^{(m)} &= \hat{\boldsymbol{\beta}}_{t|t-1}^{(m)} + \boldsymbol{G}_t^{(m)} (y_t - \hat{y}_t^{(m)}) && \text{(5d)} \\
\boldsymbol{P}_{t|t}^{(m)} &= \boldsymbol{P}_{t-1|t-1}^{(m)} - \boldsymbol{G}_t^{(m)} \mathbf{x}_{t-1}^{(m)} \boldsymbol{P}_{t-1|t-1}^{(m)}
\end{aligned}$$

where $\hat{\boldsymbol{\beta}}_{t|t-1}^{(m)} = \mathbb{E}_{t-1}(\boldsymbol{\beta}_t^{(m)})$, $\mathbb{E}_{t-1}(\cdot)$ is the expectation taken with respect to a time $t-1$ information set denoted by $\mathcal{I}_{t-1}$, and $\boldsymbol{P}_{t|t-1}^{(m)}$ is the mean square error (MSE) of $\hat{\boldsymbol{\beta}}_{t|t-1}^{(m)}$. Forecasts from model m using information set $\mathcal{I}_{t-1}$ are denoted by $\hat{y}_{t|t-1}^{(m)}$. The one-step ahead forecast error is $\hat{u}_t^{(m)}$ and its associated MSE is denoted by $F_t^{(m)}$. The $(K_m \times 1)$ vector $\boldsymbol{G}_t^{(m)}$ is the Kalman Gain. The terms $\hat{\boldsymbol{\beta}}_{t|t}^{(m)}$ and $\boldsymbol{P}_{t|t}^{(m)}$ are updated (or time t) estimates of the latent states $\boldsymbol{\beta}_t^{(m)}$ and their corresponding MSEs.

The Kalman Filter recursions in (5) are conditional on $H_t^{(m)}$ and $\boldsymbol{Q}_t^{(m)}$ (and model m). To avoid having to estimate $H_t^{(m)}$ and $\boldsymbol{Q}_t^{(m)}$, two simplifying assumptions are used in the literature. The first one, which is due to Raftery *et al.* (2010), is to replace $\boldsymbol{P}_{t|t-1}^{(m)}$ in (5a) by

$$\boldsymbol{P}_{t|t-1}^{(m)} = \frac{1}{\lambda} \boldsymbol{P}_{t-1|t-1}^{(m)}, \tag{6}$$

where $\lambda \in [0,1]$. This approximation implies that $\boldsymbol{Q}_t^{(m)} = \left(\lambda^{-1} - 1\right) \boldsymbol{P}_{t-1|t-1}^{(m)}$. In the given context, the λ parameter is commonly referred to as a *'forgetting factor'*, as it determines how many observations are effectively used for estimation.[14] The second simplifying assumption is to replace the time varying volatility $H_t^{(m)}$ by a simple exponentially weighted moving average (EWMA) estimate, that is, $H_t^{(m)}$ is constructed as:

$$H_t^{(m)} = \kappa H_{t-1}^{(m)} + (1-\kappa)\hat{u}_{t-1}^{2(m)}, \tag{7}$$

where $\kappa \in [0,1]$ is the standard EWMA smoothing parameter. Note here that an EWMA model can be thought of as a special form of a GARCH(1,1) model, ie., a restricted integrated GARCH(1,1), with the restriction being that the intercept term is fixed at 0 and that the weights on the $t-1$ volatility and squared error term sum to unity.[15]

Model averaging or selection in the DMA framework is achieved by weighting the forecasts by their respective predictive model probabilities. To clarify this, let us define $\pi_{t|t-1}^{(m)}$ to be the probability of model m given information up to time $t-1$, written as:

$$\pi_{t|t-1}^{(m)} = \Pr(\mathcal{M}_t = m|\mathcal{I}_{t-1}). \tag{8}$$

The DMA forecast of y_t, given information up to time $t-1$, denoted as $E(y_t|\mathcal{I}_{t-1})$, is then computed as:

$$\hat{y}_{t|t-1}^{(\mathrm{DMA})} = \sum_{m=1}^{M} \hat{y}_{t|t-1}^{(m)} \pi_{t|t-1}^{(m)}, \tag{9}$$

that is, as a weighted average of the forecasts from all possible models, $\{\hat{y}_{t|t-1}^{(m)}\}_{m=1}^{M}$, with the averaging weights being the predictive probabilities $\{\pi_{t|t-1}^{(m)}\}_{m=1}^{M}$.

To make the construction of the DMA forecasts in (9) feasible, model prediction and updating recursions are needed. Let $p_{jm} = \Pr(\mathcal{M}_t = m|\mathcal{M}_{t-1} = j)$ denote the (time invariant) transition probability of moving from model j at time $t-1$ to model m at time t. Also, let $f_{\mathsf{N}}^{(m)}(y_t|\mathcal{I}_{t-1})$

[14]This is also known as '*windowing*'. Intuitively, we can think of λ as a weighting function, where observations τ periods in the past receive a weight of λ^{τ}. See the discussion in Section 3.1 in Raftery *et al.* (2010) and pages $872-873$ in Koop and Korobilis (2012) for more background and intuition about the use of forgetting factors in dynamic econometric models and what it implies for the effective sample size.

[15]It is well known in the volatility literature that GARCH(1,1) models are difficult to beat in out-of-sample forecast evaluations (see, for instance, **?**). Approximating the time varying volatility by EWMA is thus unlikely to create any important loss in accuracy. We discuss later on how the κ parameter is calibrated.

denote the predictive density of y_t given model m and information up to time $t-1$. This predictive density is a Normal density evaluated at y_t with mean and variance given by $\hat{y}^{(m)}_{t|t-1}$ and $F^{(m)}_t$ as computed in (5b) and (5c), respectively. That is, $f^{(m)}_{\mathsf{N}}(y_t|\mathcal{I}_{t-1}) = \mathsf{N}(\hat{y}^{(m)}_{t|t-1}, F^{(m)}_t)$. Given an initial or prior model probability $\pi^{(m)}_{0|0}$, the model probability prediction and updating equations are then constructed as:

$$\text{Model Probability Prediction}: \quad \pi^{(m)}_{t|t-1} = \sum_{j=1}^{M} \pi^{(j)}_{t-1|t-1} p_{im} \tag{10a}$$

$$\text{Model Probability Updating}: \quad \pi^{(m)}_{t|t} = \frac{\pi^{(m)}_{t|t-1} f^{(m)}_{\mathsf{N}}(y_t|\mathcal{I}_{t-1})}{\sum_{j=1}^{M} \pi^{(j)}_{t|t-1} f^{(j)}_{\mathsf{N}}(y_t|\mathcal{I}_{t-1})}. \tag{10b}$$

A final simplification that is need to make the computation of the predictive model probabilities feasible is to approximate (10a) with

$$\pi^{(m)}_{t|t-1} = \frac{\pi^{\alpha(m)}_{t-1|t-1}}{\sum_{j=1}^{M} \pi^{\alpha(j)}_{t-1|t-1}}, \tag{11}$$

where $\alpha \in [0,1]$. The approximation in (11) has the advantage that one avoids having to specify an $M \times M$ dimensional model probability transition matrix, which would make model prediction computationally infeasible when M is large. The α parameter in (11) can again be interpreted as a '*forgetting factor*'.[16]

The implementation of the DMA procedure to forecast exchange rate returns requires the calibration of the EWMA smoothing parameter κ, as well as the two forgetting factor parameters, λ and α. We follow the guidelines provided in RiskMetrics (1996) for daily data and fix the κ parameter at 0.94.[17] For the two forgetting factors, Koop and Korobilis (2012) recommend to set the values for λ and α close to 1, so that the parameters (as well as the model probabilities) evolve reasonably gradually over time.[18] We elaborate on the choice of λ and α values in

[16]See also Section 3.2 in Raftery *et al.* (2010) and pages 874 − 875 in Koop and Korobilis (2012) for additional discussion on this.

[17]See page 97 of the documentation in RiskMetrics (1996). Note here that RiskMetrics (1996) uses λ to denote their EWMA smoothing parameter and not κ as we do.

[18]See the discussion on pages 872 − 875 in Koop and Korobilis (2012). The effective window size, ie, how much of

Section 4.

3. Data

In our empirical analysis, we use spot rates of the 6 most frequently traded currency pairs.[19] We follow standard convention in the exchange rate literature and take the US Dollar to be the home currency, so that all foreign currencies are priced in US Dollars, ie., as the US Dollar price of 1 foreign currency unit. The 6 foreign currency spot rates (denoted by S_t) are: the Euro (EUR), the Japanese Yen (JPY), the British Pound(GBP), the Australian Dollar (AUD), the Canadian Dollar (CAD) and the Swiss Franc (CHF).[20] All exchange rate data were obtained from Bloomberg. Note that we use the 5:00pm snap New York time as our '*closing*' price for the exchange rates to avoid any ambiguities related to what information was still available after the closing prices were recorded.[21] Our full data set consists of 4041 daily observations from the 4^{th} of January 1999 to the 30^{th} of June 2014. We use a 5 day working week in our analysis. Any missing data points due to, for instance, public holiday closings, are replaced by observations from the next previous available time period. This mimics the effective information flow as it is perceived by forecasting agents. We do not use interpolation methods to maintain the '*real time*' aspect of the forecasts. The choice of the sample period is driven by data availability. Data

a weight observations in the past received, is determined from $1/(1-\lambda)$ (or $1/(1-\alpha)$, respectively). Choosing values below say 0.95, would make the window narrow, so that only the very recent past would receive non-zero weights, which could result in very noisy forecasts.

[19]See page 11 in BIS (2013), which gives a list of the most heavily traded currency pairs by turnover. All data that we use are available from: www.danielbuncic.com/data/fx3data.zip.

[20]We use the academic convention to denote the spot rate as the home currency price (the US Dollar) with respect to 1 unit of foreign currency (direct quote convention from the perspective of the US Dollar). This could be made more explicit by using a USD/EUR notation (read as US Dollar price per 1 Euro) instead of just EUR. However, for reasons of compactness in notation, we will simply use EUR to mean USD/EUR. Note here that market quotes are written as EURUSD and imply 1 Euro buys a given quoted amount of US Dollars and is not to be confused with the academic notation.

[21]We use the PX_LAST entry under the Bloomberg heading, where the price data was set manually to the New York exchange values. This is important to point out, as frequently the default setting in Bloomberg is the London close. When daily data is used, there will evidently be an overlap with the information flow generated in the US and captured by the movements in the SP500, which will then carry over into the next days closing price in London. Using the 5.00pm snap New York time thus ensures that all markets have already closed on the day when the last price for the exchange rates is collected, so that this is not an issue in our analysis. Note here also that exchange rates are traded 24 hours a day, with trading at the different exchanges simply resuming once one market closes. Again, taking the 5.00pm snap, provides a clear time stamp as to what spot price was used in the return calculation.

for the Euro are available from the 4^{th} of January 1999, while yield curve data for Canada are updated with a three month lag, currently available until the end of June 2014. For Japan and Switzerland, yield curve data are only available until 21^{st} of October 2013.

To capture the influence of chartist and fundamentalist agent behaviour on exchange rates, we use two different sets of predictor variables: these are *i*) *'fundamental'* variables and *ii*) *'technical'* indicators. Fundamental variables are predictors that come from standard macroeconomic models of exchange rate determination, and include measures of aggregate output, inflation and interest rates.[22] Technical indicators are solely made up of the exchange rates' own past values. Since we are primarily interested in a *'real time'* (high frequency) forecast construction and evaluation, we use daily financial data as *'proxy variables'* for fundamentals, instead of traditional (low frequency) macroeconomic variables. Our intention here is to provide as closely as possible a *'real time'* forecast construction and evaluation scenario, which is not possible when standard low frequency data observed at monthly or quarterly intervals are used. Two well known drawbacks when using monthly (or quarterly) macroeconomic variables from aggregate accounting data are that these are released with a delay and are further subject to (potentially substantial) revisions over time as new index construction methods become available and are implemented.[23]

To avoid ambiguities with respect to data releases and revisions in our forecast evaluation, we *'extract'* information about the state of the macroeconomy from financial data, using information contained in the yield curve, stock price indices, the VIX, the TED spread, gold prices, the Baltic Dry Index (simply BDI henceforth), and the price of oil. Using financial data as a proxy for information related to macroeconomic fundamentals has the benefit of providing a clear time stamp with regards to what information was available to forecasting agents in real time.[24]

[22]See section 3 in Rossi (2013) for details regarding commonly used macroeconomic models such as the Monetary Model (MM), and models based on UIP, PPP, productivity differentials and Taylor rules.

[23]Generally, data based on aggregate accounting measures are released in intervals as new information becomes available, leading to initial, second and then final estimates. Final releases can, therefore, be up to $2-3$ months after the quarter that the data is officially recorded. For details on data revisions and a standard time line of initial, second and final releases of US GDP figures see Croushore and Stark (2001).

[24]See also, Harvey (1989), Harvey (1993) or Bodie *et al.* (2002) for examples of other studies that use financial data

We should also highlight here that we use simple returns in all our return calculations, rather than log-returns computed from differences in the log prices. That is, returns are computed as:

$$r_t = 100(P_t/P_{t-1} - 1), \tag{12}$$

where P_t is the time t price of the asset of interest. Our motivation for using (simple) returns is trivial. Due to the financial crisis period, there were at times substantial daily price variations, particularly for the BDI, gold prices, equity prices and most evidently for oil prices. For instance, on the 24^{th} of September 2001, the log-return for (WTI) oil was -16.55, while the simple return was -15.25, a difference of 1.3 percentage points. Similarly, on the 29^{th} of December 2008, the log-return and simple return were 21.28 and 23.71, respectively, a difference of around 2.4 percentage points. To be able to capture the '*true*' daily variations an investor was exposed to, we prefer to use the simple return (just return henceforth) construction as defined in (12). We should stress here also that our predictability results are not affected by the specification of the return process and hold equally well when log-returns are used.

3.1. Fundamental variables

We use three groups of financial variables as fundamental proxies to obtain information about the state of the economy — or at least as it is perceived or expected by financial market participants. These groups are: i) yield curve data, ii) stock price data, and iii) data related to risk aversion and global trade activity.

3.1.1. Information in the yield curve

The use of yield curve data is motivated by the findings in Bekaert and Hodrick (1992) and Clarida *et al.* (2003), who show that the information content in the yield curve is valuable for exchange rate forecasting.[25] Moreover, in the context of macro-finance models, Diebold *et al.*

as proxies for macroeconomic variables.

[25] In a related context, high frequency yield curve data has been studied in Gürkaynak *et al.* (2005) and Brand *et al.* (2010) to asses the effect of central bank communication on various asset prices, including equities and exchange rates.

(2006) and Rudebusch and Wu (2008) have documented that the empirically derived level and slope factors are *'strongly'* correlated with inflation and economic activity.[26]

To capture the information in the yield curve, we construct level, slope and curvature factors, denoted by $\mathcal{L}_t$, $\mathcal{S}_t$ and $\mathcal{C}_t$, using daily data on zero coupon yields. We follow Diebold *et al.* (2006) and compute the three *'empirical'* factors using linear combinations of yields of various maturities.[27] The level, slope and curvature factors are computed as:

$$\text{Level}: \ \mathcal{L}_t = \left(y_t^{(3)} + y_t^{(24)} + y_t^{(120)}\right)/3 \tag{13a}$$

$$\text{Slope}: \ \mathcal{S}_t = \left(y_t^{(3)} - y_t^{(120)}\right) \tag{13b}$$

$$\text{Curvature}: \ \mathcal{C}_t = \left(2y_t^{(24)} - y_t^{(3)} - y_t^{(120)}\right), \tag{13c}$$

where $y_t^{(\tau)}$ is the time t yield of a zero-coupon bond with maturity τ (measured in months). Zero coupon data for the US are taken from the well known and widely used Gürkaynak *et al.* (2007) database. For Australia, Canada and the UK, they are taken from the websites of the Reserve Bank of Australia (RBA), the Bank of Canada (BoC), and the Bank of England (BoE). For the euro area, the available yield curve data from the European Central Bank (ECB) only go back to the beginning of September 2004. To extend the data to the beginning of January 1999, we use yield curve factors from the Bundesbank before September 2004.[28] Due to the lack of publicly available daily data for Switzerland and Japan, we use the (daily) Nelson-Siegel-Svensson parameter estimates from Malkhozov *et al.* (2014) to construct the yield curve factors

[26] More specifically, Diebold *et al.* (2006) show that their empirical level and slope factors have correlations of 43% and 39% with (year on year price deflator) inflation and capacity utilisation respectively (see page 319). Similarly, in a New Keynesian macro-finance model, Rudebusch and Wu (2008) find that their (macro-finance) level and slope factors have a 73% and 66% correlation with 1-year expected inflation and output (see page 916).

[27] Since it is common to use the Nelson-Siegel-Svensson approach to construct the zero coupon yields, as is done, for instance, by Gürkaynak *et al.* (2007), one could also use the slope coefficients fitted from the cross-sectional regression of the yields (ie., the $\beta_i, \forall i = 0, 1, 2$ estimates). Nevertheless, as is evident from the estimates that are provided in the Excel file provided by Gürkaynak *et al.* (2007) at http://www.federalreserve.gov/pubs/feds/2006, there can be considerable variation over time. We therefore prefer to construct the level, slope and curvature factors from the actual yield data.

[28] Note that both, the ECB and the Bundesbank, use the parametric approach of Svensson (1994) to construct zero coupon yields, so the methods of construction are consistent, despite the parameters being calibrated on two different sets of bonds. For the euro area, we use the Svensson (1994) parameter estimates reported under the *'all issuers whose rating is triple A'* heading.

for these two countries.[29] The sample period that is covered by Malkhozov *et al.* (2014) nevertheless ends on 21st of October 2013, thereby shortening the available out-of-sample evaluation period for the Yen and Swiss Franc somewhat.

To provide information about each countries perceived macroeconomic fundamentals as capture by the yield curve factors relative to the US economy, we construct differences between the level, slope and curvature factors of the US and the foreign currency of interest. These are denoted by $\mathbf{x}_t^{\mathrm{LSC},i}$ and computed as:

$$\mathbf{x}_t^{\mathrm{LSC},i} = \Big[(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^i),\ (\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^i),\ (\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^i)\Big], \tag{14}$$

where $\mathcal{L}_t^{\mathrm{US}}$ $(\mathcal{L}_t^i)$, $\mathcal{S}_t^{\mathrm{US}}$ $(\mathcal{S}_t^i)$, and $\mathcal{C}_t^{\mathrm{US}}$ $(\mathcal{C}_t^i)$ are level, slope and curvature factors for the US (the i^{th} foreign currency), respectively, with $i = \{\mathrm{EU, JP, GB, AU, CA, CH}\}$ being a country index for the exchange rates of interest. Due to the high persistence in the yield curve factors $\mathbf{x}_t^{\mathrm{LSC},i}$, we use the (time) difference of the yield curve factors denoted by $\Delta\mathbf{x}_t^{\mathrm{LSC},i}$, with Δ being the difference operator, as the predictor variables in the forecast evaluation.[30]

3.1.2. Information in stock prices

We add stock price indices to the set of predictor variables to complement the information on macroeconomic fundamentals as contained in the yield curve. The usefulness of the information content embedded in US (as well as other countries') stock returns for the purpose of forecasting the equity premium has recently been demonstrated by Rapach *et al.* (2013).[31] We use the SP500, as well as each individual country's head line stock price index, in the set of fundamental predictors. The headline indices are: the Nikkei225 for Japan, the FTSE100 for

[29] We thank Andrea Vedolin for making these parameter estimates available to use. For more details on the construction of the factors and the bond data that was used, we refer to Malkhozov *et al.* (2014).

[30] Level, slope and curvature factors are known to be highly persistent, especially at daily frequencies. For the US, for instance, the first order autocorrelations are, 0.9994, 0.9986, and 0.9974 for $\mathcal{L}_t^{\mathrm{US}}, \mathcal{S}_t^{\mathrm{US}}$ and $\mathcal{C}_t^{\mathrm{US}}$, respectively. Computing yield curve factors relative to the US ones remain highly persistent. Using euro area factors, this difference has autocorrelations of 0.9984, 0.9962, and 0.9914 for $(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^{\mathrm{EU}}), (\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^{\mathrm{EU}})$ and $(\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^{\mathrm{EU}})$, respectively. To avoid this high persistence, we prefer to work with the (time) differenced series for $\mathbf{x}_t^{\mathrm{LSC}}$, that is, $\Delta\mathbf{x}_t^{\mathrm{LSC}}$.

[31] Rapach *et al.* (2013) show that lagged US returns have significant predictive power to forecast equity premia in numerous industrialised countries. The economic intuition behind this finding is given in relation to the US being an information originator (see pages $1635 - 1636$ in Rapach *et al.* (2013) for a detailed explanation).

the UK, the SPI for Switzerland, the SPTSX for Canada and the All Ordinaries for Australia.[32] For the euro area, it would seem natural to opt for the EURO STOXX 50 as a representative stock price index. Nevertheless, the key headline index from the view point of the financial media still seems to be the DAX30, not only for Germany, but for the euro area as a whole. The DAX30 is also a more liquid market index. It has an approximately 50% higher trading volume (3 months average) than the EURO STOXX 50. For this reason, we prefer to use the DAX30 as the headline index for the euro area.

We construct returns of the stock price indices of the SP500 and each country's headline index to be used as predictor variables. That is, the stock (or equity) price predictor set (denoted by $\mathbf{x}_t^{\mathrm{EQT},i}$) consists of:

$$\mathbf{x}_t^{\mathrm{EQT},i} = \left[r_t^{\mathrm{EQT,US}},\ r_t^{\mathrm{EQT},i} \right], \tag{15}$$

where $r_t^{\mathrm{EQT,US}}$ and $r_t^{\mathrm{EQT},i}$ denote the return on the SP500 US stock price index and the return on the i^{th} foreign equity market corresponding to the currency of interest, respectively, computed as defined in (12) for each equity price index. The i superscript here is again used to denote the foreign headline equity price index corresponding to the exchange rate of interest, ie., $i = \{\text{DAX30, Nikkei225, FTSE100, AllOrds, SPTX, SPI}\}$.

3.1.3. Risk aversion measures and global trade activity

In addition to the yield curve and stock price data, we also include variables that are meant to capture risk aversion and global trade activity in the set of fundamental predictors.

We use the VIX index and the TED spread to provide us with a *'sense of risk aversion'* in the market. The VIX measures the volatility implied by option prices on the SP500 and thus reflects investors' expectations about stock market volatility over the next month.[33] The TED spread is calculated as the difference between the 3 month LIBOR rate (US dollar base) and the 3 month Treasury Bill rate and measures the perceived credit risk in the US economy. A higher value

[32] For Australia, we prefer to use the All Ordinaries index over the SP/ASX200 because of its longer data history, but also because it constitutes a broader index, containing 500 of the largest stocks as opposed to only 200 as the SP/ASX200.

[33] The VIX is computed as the weighted average of the implied volatilities of options on the SP500 index for a wide range of strikes and mainly first and second month expirations

in the VIX and/or the TED spread is generally taken as an indication of market participants expecting an overall negative economic or financial outlook, and hence an increased (global) aversion to risk.[34] Brunnermeier *et al.* (2009) have shown that the VIX and the TED spread predict higher returns in carry trade strategies which are widely used by foreign exchange traders.

We also include gold as a viable predictor variable. The motivation for this is twofold. First, gold is considered to be a *'safe haven'* asset and hence constitutes a complement to the VIX and the TED spread indicators of risk aversion in financial markets. Gold is further regarded to be a hedge against inflation, deflation, as well as general uncertainties related to economic, financial and political instabilities. Second, together with other precious metals such as platinum and silver, gold is also commonly held in investment portfolios that are diversified over equities, bonds and exchange rates. Gold can thus be seen as a natural portfolio complement to foreign currency holdings in an investment portfolio. We expect, therefore, movements in gold prices to be informative for exchange rate forecasting, particularly since the financial crisis in 2008.

As a proxy for global trade flows as well as supply and demand trends in production of finished goods and raw materials, we include the Baltic Dry Index (BDI) and crude oil prices as fundamental predictor variables. The BDI is a composite index of the Baltic Capesize, Panamax, Handysize and Supramax indices. This index is designed as the successor to the Baltic Freight Index. The BDI is frequently viewed as a leading indicator of future global trade demand and economic growth, as the goods that are shipped are raw materials and thus give an indication of the demand for primary production inputs (see, for instance, Baumeister and Kilian (2014) who also use the BDI to forecast oil prices).

The rational for using crude oil prices in the set of predictors is due to oil still being one of the most widely used sources of energy (see for instance, among many other studies, the evidence reported in Lardic and Mignon (2008) and He *et al.* (2010)). Moreover, there is a widely

[34] The VIX and the TED spread are widely regarded as measures of the *'global appetite for risk'*. This is not only the case for equity markets and equity options markets, but also for corporate credit markets and foreign exchange markets (see for instance the evidence reported in Collin-Dufresne *et al.*, 2001). Also, Pan and Singleton (2008) find that the VIX in particular is strongly related to the variation in risk premiums in sovereign credit default swaps.

held view that unexpected increases in the price of oil can cause recessions in many oil importing countries (see Kilian (2008), Hamilton (2009) and others). High oil prices are often also linked to periods of higher inflation, thereby directly affecting central bank policy and thus the setting of interest rates (Bhar and Mallik, 2013). Lastly, oil prices, in conjunction with US Energy Information Administration (EIA) inventories are closely monitored by financial market participants and reported in the financial press. These are taken to be early indicators of changes in production and manufacturing demand.

The risk and global trade activity predictor set, which we denote by $\mathbf{x}_t^{\text{RISK/ACTIV}}$, includes the following variables:

$$\mathbf{x}_t^{\text{RISK/ACTIV}} = \left[\Delta\text{VIX}_t,\ \Delta\text{TED}_t,\ r_t^{GOLD}, r_t^{BDI},\ r_t^{OIL}\right], \tag{16}$$

where ΔVIX_t, and ΔTED_t denote the (time) difference in the series of the CBEO Volatility Index and the TED spread, and r_t^l, $\forall l = \{\text{Gold, BDI, Oil}\}$ are the returns from investing in gold, the Baltic Dry Index and oil, again, as defined in (12). Note here that we use the differences in the VIX and TED spread series. We could have also used the level series instead. Nevertheless, since both series are once again highly persistent, we have opted for the difference specification as used in Brunnermeier *et al.* (2009) as well.[35]

All fundamental predictors which are used by fundamentalist agents to form their forecasts of currency i at time t are collected in the (10×1) dimensional vector:

$$\left[\mathbf{x}_t^{\text{LSC},i},\ \mathbf{x}_t^{\text{EQT},i},\ \mathbf{x}_t^{\text{RISK/ACTIV}}\right], \tag{17}$$

where i denotes the foreign currency of interest.

[35] Evidently there will be some loss of *'information'* when considering changes only, that is, irrespective of the level of either the VIX or the TED spread series. Clearly, a given unit change at a high VIX or TED spread level, will have a different impact than the same change at a much lower level. Rather than just using the level or the differenced series, one could also take some interactions that account for the changes as well as the level of the series. Nonetheless, to avoid specification searches that yield the best *'statistical results'*, we do not do this here and use simply the differenced form to remove the heavy persistence in the series.

3.2. Technical variables

To enhance our *'macroeconomic fundamentals'* information set, we construct various technical indicators. Technical analysis is a widespread method employed by market participants to forecast largely short term movements in asset prices. Neely *et al.* (2014) have recently successfully used technical indicator variables as predictors to forecast the equity risk premium.

Technical analysis involves using charts of financial asset price movements combined with additional descriptive statistics to infer the likely course of future prices and hence to form trading strategies. Often, chartists use trends and patterns in general to identify broad ranges within which exchange rates or asset prices are expected to trade. Also they employ mechanical indicators, which may be trend-following (based on moving-averages) or non-trend following indicators (such as reversal indicators) with the assumption that there is a tendency for markets to correct. In practice, technical analysis is a combination of pattern and trend recognition, along with information from basic statistical indicators (see also Sarno and Taylor (2003) for more details on technical trading rules and its use in foreign exchange markets).[36]

We follow the approach in Neely *et al.* (2014) and construct various technical indicators to be used as predictors in our chartist model. These technical indicators are grouped into the following blocks: (i) Moving Average (MA) rules, (ii) Momentum indicators, and (iii) indicators based on the Relative Strength Index (RSI).[37]

[36]There is strong evidence from survey data of high use of technical analysis at short time horizons (intraday to one week), while there is a skew towards macroeconomic fundamentals for longer time horizons (see, Allen and Taylor (1990), Taylor and Allen (1992), Menkhoff (1998), Lui and Mole (1999), among many others). Although there seem to exist some market participants that solely rely on technical analysis, most dealers perceive technical analysis and fundamental analysis to be complementary approaches. Also a significant proportion of market participants view technical analysis as self-fulfilling. Allen and Taylor (1990) report that some chartists were able to outperform a random walk model in one and four week ahead forecasts over a ten-month sample period. Furthermore, Allen and Taylor (1990) report a significant degree of heterogeneity among chartists' forecasts, that is, they see the same asset price but different signals, and interpret them in a different manner at the same point in time. Lui and Mole (1999) find moving averages and other trend-following systems to be the most applied technical methods. Bilson (1990) shows that chartists employing oscillators such as relative strength indices, that indicate overbought and oversold conditions, are able to impart nonlinearity into exchange rate movements. Small exchange rate changes which do not trigger the oscillator will tend to be positively correlated because of the effect of trend-following trading programs, whereas larger movements, which trigger an oscillator, will be negatively correlated. The fact that a lot of market agents apply technical analysis implies that it should not be dismissed in a set of variables used for forecasting purpose.

[37]The first two indicators are the same as those use in Neely *et al.* (2014). We add the RSI indicator, as it is another widely used technical indicator that measures the level of *'overreaction"*, that is, overbought or oversold conditions, in asset prices.

3.2.1. Moving Average Rules

Moving average (MA) cross-over rules are among the most popular and common trading rules discussed in the technical analysis literature (Sullivan *et al.*, 1999, p. 1656). The standard cross-over rule, as outlined in Gartley (1935), is that the down penetration of the MA by the price is regarded as a sell-signal and the upside penetration as a buy signal. There are various modifications to this rule. Buy and sell signals can be generated by crossovers of slow and fast MAs, where a slow MA is computed over a longer number of days than a fast MA.

Formally, the moving average of the exchange rate computed over the last n daily closing prices is defined as:

$$\mathrm{MA}_t^{(n)} = n^{-1} \sum_{i=t-n+1}^{t} S_i, \;\; \forall\, t \geq n \in \mathbb{Z}, \tag{18}$$

where S_t is the daily spot closing price and n is the number of days that are averaged over. We consider the simplest and most widely used long-term cross-over of the 200 day moving average $\mathrm{MA}_t^{(200)}$ and the spot price S_t.[38]

The $\mathrm{MA}_t^{(200)}$ is commonly viewed as a long-term trend indicator, with the indicator generating a broad buy signal as long as $S_t > \mathrm{MA}_t^{(200)}$, while the penetration of the $\mathrm{MA}_t^{(200)}$ by S_t from above reverses the signal to a sell indicator. Formally, we define the $S_t > \mathrm{MA}_t^{(200)}$ buy indicator as:

$$\mathrm{IMA}_t^{(200)} = \begin{cases} 1 & \text{if } S_t > \mathrm{MA}_t^{(200)} \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

[38]Note that there are many other viable cross-over candidates involving cross-over rules of slow and fast moving MAs. These are generally the 50 and 100 day cross-overs with the 200 day MA. Nevertheless, to avoid any ambiguities related to *'searching over the best cross-over rule'* issues, we stick to the simplest and most widely used long-term cross-over indicator, the crossing of the spot price S_t and the 200 day MA (see here also the relatively recent post on www.marketwatch.com with the title: What breaking the 200-day moving average for stocks really means for recent media coverage (article was published October 14, 2014) based on the spot price breaking through the 200 day MA.

3.2.2. Momentum Indicators

As an alternative to the moving average indicator, we also include a simple momentum indicator in the set of technical predictor variables. Momentum indicators are meant to capture the sentiment or trend following component in exchange rates, that is, the strategy to buy a currency if it had a positive return over the last n periods, and sell a currency if had a negative return. We use a time period of $n = 130$ days (6 months) to measure momentum, and define the 130 day momentum indicator as:

$$\text{IMOM}_t^{(130)} = \begin{cases} 1 & \text{if } S_t > S_{t-130} \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

The choice for 130 trading days (which corresponds to approximately half a year when 260 annual trading days are assumed) is mainly driven by a trade-off between the ability to capture known "*long-swings*" in exchange rate data and to adapt quickly to recent changes. In the equity premium forecasting literature, it seems to be more common to use 9 months or 12 months horizons to compute the momentum indicator (see for instance page 4 in Neely *et al.* (2014)). Nevertheless, the choice of using 6 months rather than 12 months returns to compute the momentum indicator does not have any important implications for our predictability results.[39]

3.2.3. Relative Strength Index

We use the 14 day relative strength indices, denoted by $\text{RSI}_t^{(14)}$, in addition to the moving average and momentum indicators in the set of technical indicators. The RSI, as developed by Wilder (1978), measures the velocity of a security's price movement to identify overbought and oversold conditions. There exists recent empirical evidence illustrating the success of RSI based trading strategies. For instance, Chong and Ng (2008) use RSI based trading rules on the London Stock Exchange FT30 Index to analyze if these are profitable. Their conclusion is that

[39] Also, Neely *et al.* (2014) use volume data as a technical indicator. We do not do this here largely due to data availability. Volume data is much more difficult to get hold of, as exchange rates are still traded to a large extent over-the-counter.

an RSI based trading strategy is able to out-perform a simple buy-and-hold strategy. Similarly, Rodríguez-González *et al.* (2010) employ RSIs in a neural network context to predict individual stocks and are able to predict more than 50% of directions of change.

For a general n, the RSI is constructed as:

$$\mathrm{RSI}_t^{(n)} = 100 - \frac{100}{1 + \frac{\mathrm{MA}_t^{(n)}(dc_t)}{\mathrm{MA}_t^{(n)}(uc_t)}} \tag{21}$$

where $\mathrm{MA}_t^{(n)}(\xi_t)$ denotes the $n-$period MA filter in (18) applied to variables ξ_t, and uc_t (dc_t) are upclose (downclose) measures defined as:

$$uc_t = \begin{cases} \Delta S_t & \text{if } \Delta S_t > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad dc_t = \begin{cases} -\Delta S_t & \text{if } \Delta S_t < 0 \\ 0 & \text{otherwise,} \end{cases} \tag{22}$$

with $\Delta S_t = S_t - S_{t-1}$ being the difference of the spot price of the exchange rate and n again the number of days over which the uc_t and dc_t are averaged over. Note that RSIs are, by construction, an index over the 0 to 100 range.[40]

To account for possible (traditional) time series dynamics in the returns, we also add lagged values of the returns to the set of technical predictors. The full set of technical predictor variables that is used by chartists to construct forecasts for the i^{th} foreign currency is composed of the following 4 variables:[41]

$$\left[r_t^i,\ \mathrm{IMA}_t^{(200)},\ \mathrm{IMOM}_t^{(130)},\ \mathrm{RSI}_t^{(14)} \right]. \tag{23}$$

[40]When working with stock price data, a stock is considered to be overbought when its RSI is above 70 and as oversold when is RSI is below 30. The choice of $n = 14$ is due to this being the most prominent value used among technical analysts, and in many software programs is the default setting. Our results do not change in any important way if we use $n = 20$ instead, which is another popular setting.

[41]Note here, that, for simplicity of notation, we do not include i index counters on the technical indicators, but it should be clear that these are computed for the currency of interest.

3.3. Summary statistics and visual overview

In this section, we briefly describe some of the basic features of the exchange rate data and the two different sets of predictor variables that we use in our forecast evaluation. Note that the discussion here is not meant to be exhaustive, but rather complementary to the tables and figures that we provide to summarise this information.

← Figure 1 about here

In Figure 1 we show plots of the 6 different currencies that are used in the forecast evaluation exercise over the full sample period from January 4, 1999 to June 30, 2014. The left column in Figure 1 shows the (raw) US Dollar price of one foreign currency unit, and the right column shows the daily returns of the series. As a reminder, an upward movement in these 6 series indicates that the respective currency has appreciated against the US Dollar (the US Dollar price of the foreign currency has risen), while a downward movement suggests a depreciation. There are a number of interesting visual features that are evident from the plots in Figure 1. First, notice how all 6 exchange rates show a general upward trend, suggesting that the series have appreciated over the last 15 years against the US Dollar. This trend is much weaker for the British Pound series from mid September 2008 onwards. Second, the Lehman Brothers collapse in September 2008 had a rather profound effect on the British Pound, the Australian Dollar, the Canadian Dollar and the Euro, resulting in depreciations of approximately 22%, 18%, 17%, and 10%, respectively, from September 1, 2008 to March 1, 2009.[42] These four currencies thus behaved inline with what would be expected from an *'investment currency'*, where high levels of risk aversion lead to sell-offs in such a particular asset class.

Over the same time frame, the Japanese Yen appreciated by nearly 11%, while the Swiss Franc remained rather stable, depreciating only marginally by 2%. It is interesting to see here that the Japanese Yen behaved in accordance with its widely perceived *'safe haven'* status, that is, providing financial refuge during times of high risk aversion, while the Swiss Franc, also known as a *'safe haven'* currency, was largely unaffected. The Swiss Franc's *'safe haven'* status did not come to bear any significant importance until the first set of problems began which

[42] For all 4 currencies, the depreciation relative to the US Dollar already started somewhat earlier, nevertheless, it is clear from the plots that from September 2008 the drops in the currencies amplified substantially.

eventually lead to the European sovereign bond crisis. The strongest Swiss Franc appreciation was realised over the July 2010 to August 9, 2011 period, where the currency surged over 41% against the US Dollar.[43] It is interesting to point out here that, although the European sovereign bond crisis appears to have been one of the key drivers of the *'save haven'* effect of the Swiss Franc, the strong appreciation in the Franc was not matched by an equal depreciation in the Euro against the US Dollar. The Euro, in fact, appreciated by nearly 17% over the same time period, highlighting that there must have been other additional factors that contributed to the strong appreciation of the Swiss Franc. The fact that the Euro did not depreciate against the US Dollar as a result of the European sovereign bond crisis could partly be due to the second round of the quantitative easing program of the Federal Reserve in the US being implemented.

In the right column of Figure 1, the time series evolution of the return series is plotted. As was the case in the levels plot of the 6 series, the homogenous response to the Lehman Brothers collapse in September 2008 is also clearly visible in all 6 return series of the currencies, with the response of the Swiss Franc, nevertheless, being somewhat weaker than for the other currencies. The Australian Dollar is the most volatile currency, with daily returns in the AUD swinging between -8 to 8 percent throughout most of October 2008. The intervention by the Swiss national bank via the imposition of the cap on the CHF/EUR rate at 1.20 is the most outstanding event impacting on the 6 return series after the Lehman Brothers collapse.

In Table 1 we show summary statistics of the 6 exchange rate returns to be forecasted (top 6 rows), as well as summary statistics of the fundamental predictor variables that are used by fundamentalist agents over the full sample period from January 4, 1999 to June 30, 2014.[44] The fundamental predictor variables are arranged coherently in three separate blocks to match the description of the variables in the text. Looking over the summary statistics of the 6 exchange rates, a number of the stylised facts which were visible in Figure 1 are also evident in the sum-

[43]This surge may have been triggered by a number of events, starting with Greece needing "*official*" financial assistance in May 2010, followed by Ireland's bailout in November 2010 and with Portugal following in May 2011 (see Lane, 2012, page 56). Note here also that the Swiss national bank imposed the cap on the CHF/EUR rate at 1.20 on September 6, 2011, but speculation about the implementation of the cap had already been circulating for weeks beforehand, so that the peak in fact occurred on August 9.

[44]For Japan and Switzerland, the level slope and curvature factors only go up to October 21, 2013.

mary statistics. First, the mean return on the 6 exchange rates series is positive, confirming the positive trend in the level series and the overall depreciation of the US Dollar against these 6 currencies. Second, the Australian Dollar is the most volatile series, with a (daily) standard deviation of 0.84 percent, followed by the Swiss Franc with a standard deviation of 0.69 percent. These two currencies have further the *'heaviest tails'*, as is evident from their kurtosis statistics being well above 10.[45] What is interesting to point out from the summary statistics of the exchange rate returns are the rather sizable negative first order autocorrelations (henceforth ACF(1)) for the returns on the Yen, Canadian and Australian Dollars, as well as the Swiss Franc, as shown in the last column of Table 1.

Looking over the summary statistics of the fundamental predictor variables that are reported below the solid line separating the exchange return series from the predictors in Table 1, it is evident that the most volatile predictors are oil returns, BDI returns, as well as DAX30 and Nickkei225 equity returns. All equity returns as well as returns from investing in gold, oil and the BDI had positive means, raging from values as low as 0.01% for the return on the FTSE100 up to 0.08% for the return on oil. The return from investing in oil was not only the most volatile series, but it also had the single largest positive daily return of nearly 24%. From the first order ACFs in the last column of Table 1 it can be seen that all but one equity return series are negatively correlated, with the SP500 returns showing the strongest negative autocorrelation of -0.08. The single equity series with a positive ACF is the return on the Swiss SPI. By far the strongest first order autocorrelated predictor variable is the return on the BDI with an ACF(1) of 0.80.[46]

For reasons of completeness, we show the equivalent summary statistics corresponding to the technical indicators as used by chartist trading agents in Table 2. We group the three technical indicators of the 6 currencies into blocks for ease of readability. Note initially from the summary statistics that the moving average and momentum technical indicators, as defined in

[45]Note that this here is kurtosis and not excess kurtosis, so should be measured against a benchmark of 3. Nevertheless, this is still pretty high when compared to the other 4 currencies.

[46]The autocorrelation and partial-autocorrelation structure of the BDI returns (not shown here) in fact display the properties of an ARMA$(1, 2)$ model.

(19) and (20), are binary, so that the means measure the proportion of time that a buy signal was generated by the technicals. From these means one can notice that all indicators give readings of less than 0.5, suggesting that buy signals were generated less than 50% of the time over the approximately 15 years of data that are used in the forecast evaluation. The highest proportion of moving average based buy signals is generated for the British Pound, with the lowest being for the Australian Dollar. For the momentum based indicator, the largest proportion of buy signals is for the Yen, while the lowest is for the Swiss Franc. Overall, we can also notice from the first three ACFs and PACFs which are reported in the last six columns of Table 2 that the binary indicators for the moving average and momentum rules generate a fair degree of persistence in the predictor variables, with not only sizeable ACFs, but also PACFs.

The relative strength index based regressors, which are continuous but bounded series in the $[0, 100]$ interval, are broadly centered at an RSI value of close to 50 for all 6 exchange rate series. Skewness and Kurtosis statistics indicate that the RSI based technical indicators are fairly symmetric without any showing of heavy tails. The ACFs and PACFs for the RSIs also indicate a noticeable degree of first order persistence in the series, similar to the persistence that an AR(1) process would generate.

4. Forecast construction and evaluation

We now describe in detail how the forecasts are constructed and how the evaluation is carried out. Since we are primarily interested in the real time predictive performance of the model, we implement an out-of-sample forecast evaluation of the model. Also, before we outline in detail the statistical criteria that we utilise to assess the performance of the model, and before the results of the forecast evaluation are reported, we initially describe the prediction setting that we use in our evaluation.

4.1. Prediction setting

As outlined in Section 2, we implement an agent based model averaging/combination approach to forecast the returns of our 6 exchange rates of interest. The agent based model combined predictions (henceforth, simply MC predictions, which we denote by $\hat{r}^{\mathsf{MC}}_{t+1|t}$) are computed as:

$$\hat{r}^{\mathsf{MC}}_{t+1|t} = (1 - \hat{\omega}_t)\hat{r}^{\mathsf{F}}_{t+1|t} + \hat{\omega}_t \hat{r}^{\mathsf{C}}_{t+1|t}, \tag{24}$$

where $\hat{r}^{\mathsf{F}}_{t+1|t}$ and $\hat{r}^{\mathsf{C}}_{t+1|t}$ denote, respectively, the individual $1-$step ahead forecasts constructed by fundamentalist and chartist agents at time t. The (fitted) weights $\hat{\omega}_t$ in (24) are obtained by means of a Granger and Ramanathan (1984) forecast combination regression of the form:

$$r_t = (1 - \omega_t)\hat{r}^{\mathsf{F}}_{t|t-1} + \omega_t \hat{r}^{\mathsf{C}}_{t|t-1} + \nu_t \tag{25}$$

which, for estimation purposes, can be conveniently re-written in the equivalent form:

$$(r_t - \hat{r}^{\mathsf{F}}_{t|t-1}) = \omega_t(\hat{r}^{\mathsf{C}}_{t|t-1} - \hat{r}^{\mathsf{F}}_{t|t-1}) + \nu_t, \tag{26}$$

where the weights ω_t in (26) are restricted to be in the $[0, 1]$ interval.[47]

Two features are evident from the specification of the model averaged predictions defined in (24) and how the combination weights are obtained in (26). First, the time t estimate of ω_t is used in the averaged forecast construction in (24). This implies that the best forecast of the averaging weight ω_t one period into the future is its time t estimate, that is, $E_t(\omega_{t+1}) = \hat{\omega}_t$. The weights are thus assumed to evolve as a random walk process. Second, individual return forecasts from the fundamentalist and chartist agents are needed. We describe below in the next section how these are computed using the DMA methodology that was outlined in Section 2.1.

[47] Imposing the constrain $\omega_{t,h} \in [0, 1]$ can be easily implemented in the simplest from via a grid search, or via a standard constrained least squares (or quadratic programming) algorithm. We use the `lsqlin` command in Matlab to enforce the $[0, 1]$ interval on $\omega_{t,h}$.

4.1.1. Computing individual fundamentalist and chartist forecast

To avoid unnecessary clutter in the notation that follows, we drop the i term that indexes the 6 different exchange rates that we model and describe how fundamentalist and chartist agents form their forecasts for a generic investment currency. Also, let $\mathsf{a} = \{\mathsf{F}, \mathsf{C}\}$ denote the agent type for which we construct the forecast, ie., a fundamentalist or chartist agent type. Following the general description of the DMA framework in Section 2.1, the forecasting model for (the i^{th}) exchange rate return of agent a takes the form

$$r^{\mathsf{a}}_{t+1} = \mathbf{x}^{\mathsf{a}(m)}_t \boldsymbol{\beta}^{\mathsf{a}(m)}_{t+1} + u^{\mathsf{a}(m)}_{t+1} \tag{27a}$$

$$\boldsymbol{\beta}^{\mathsf{a}(m)}_{t+1} = \boldsymbol{\beta}^{\mathsf{a}(m)}_t + \boldsymbol{\epsilon}^{\mathsf{a}(m)}_{t+1}, \tag{27b}$$

where $m = 1, \ldots, M$ denotes the model index, r_{t+1} is the one-period holding return of the currency of interest, and the full predictor set $\mathbf{x}^{\mathsf{a}}_t, \forall \mathsf{a} = \{\mathsf{F}, \mathsf{C}\}$ is, following (17) and (23), defined as:

$$\mathbf{x}^{\mathsf{F}}_t = \left[1, \mathbf{x}^{\text{LSC}}_t, \mathbf{x}^{\text{EQT}}_t, \mathbf{x}^{\text{RISK/ACTIV}}_t\right] \tag{28}$$

for fundamentalists and

$$\mathbf{x}^{\mathsf{C}}_t = \left[1, r_t, \text{IMA}^{(200)}_t, \text{IMOM}^{(130)}_t, \text{RSI}^{(14)}_t\right] \tag{29}$$

for chartist agents. In (28) and (29), the intercept term is denoted by 1, and the lagged exchange rate return r_t in (29) allows for the possibility of AR type dynamics in the returns.[48] The number of predictor variables (excluding the intercept term) is 10 and 4, so that a total of $2^{10} = 1024$ and $= 2^4 = 16$ models are available, at each point in time to fundamentalist and chartist agents, respectively.[49]

[48] It is clear that there is only very weak evidence of any autocorrelation in the return series from the summary statistics that we report in Table 1. Nevertheless, these summary statistics are an unconditional measure computed over the full sample period. There may be times when there is an (absolute value) increase in return autocorrelation for different periods of time. To allow for this possibility, we decided to include lagged returns in the chartist predictor set.

[49] Note that the null or base model is the one with only an intercept term in it.

To compute the return forecasts for the exchange rates from the recursions in (27), time t filtered estimates $\hat{\boldsymbol{\beta}}^{\mathsf{a}(m)}_{t|t}$ are needed to construct the optimal forecast of $\boldsymbol{\beta}^{\mathsf{a}(m)}_{t+1}$. Given our random walk specification of the state dynamics in (3b), this forecast is given by $\hat{\boldsymbol{\beta}}^{\mathsf{a}(m)}_{t|t}$, for all $m = 1, \ldots, M$. The sequence of $\hat{\boldsymbol{\beta}}^{\mathsf{a}(m)}_{t|t}$ is obtained from the Kalman filter recursions outlined in (5). To implement the Kalman filter, we need to specify initial values. We follow Koop and Korobilis (2012) and use a diffuse prior for $\hat{\boldsymbol{\beta}}^{\mathsf{a}(m)}_{0|0} \sim \mathsf{MN}(\mathbf{0}_{K_m}, 100\boldsymbol{I}_{K_m})$, where $\mathbf{0}_{K_m}$ is a $(K_m \times 1)$ dimensional vector of zeros and $\boldsymbol{I}_{K_m}$ is $(K_m \times K_m)$ dimensional identity matrix. The model updating probabilities for agent type a in (10b) are initialised with an uninformative prior $\pi^{(m)}_{0|0} = \frac{1}{M}$, so that all models are assumed to be equally likely. The α and λ parameters are set to $(0.99, 0.999)$.[50] The κ term in the EWMA specification is fixed at 0.94, in line with current RiskMetrics (1996) recommendations for daily data.[51]

Given the model updating probabilities $\pi^{\mathsf{a}(m)}_{t|t}$ and $\alpha = 0.99$, forecasts of the model probabilities are computed as $\pi^{\mathsf{a}(m)}_{t+1|t} = \pi^{\mathsf{a}\alpha(m)}_{t|t} / \sum_{j=1}^{M} \pi^{\mathsf{a}\alpha(j)}_{t|t}$, yielding the DMA based forecast of agent type a of the exchange rate return of interest as:

$$\hat{r}^{\mathsf{a}}_{t+1|t} = \sum_{m=1}^{M} \mathbf{x}^{\mathsf{a}(m)}_{t} \hat{\boldsymbol{\beta}}^{\mathsf{a}(m)}_{t|t} \pi^{\mathsf{a}(m)}_{t+1|t} \tag{30}$$

for $\mathsf{a} = \{\mathsf{F}, \mathsf{C}\}$, that is, the forecasts of fundamentalist or chartist agent type.

4.1.2. Fitting and evaluation periods

Our entire available data set consists of $T = 4039$ $(T = 3860)$ observations, covering the period from January 4, 1999 to June 30, 2014 for the EUR, GBP, AUD and CAD, and for JPY and

[50] Koop and Korobilis (2012) and Raftery *et al.* (2010) note that it is common to choose a λ value near 1 for a gradual evolution of the coefficients in the models. Since we are using daily data, we would like to avoid too much variation in the coefficients by using too low values for λ. We therefore set λ at 0.999. To put this in the context of the comparison that Koop and Korobilis (2012) carry out on page 872 in terms of how much weight observations a fixed period in the past receive, with $\lambda = 0.99$, observations one year ago (that is, 260 days) receive a weight of only 7.33%. This is much too low and would imply substantial variability in the coefficients over time. With $\lambda = 0.999$ this weight is 0.7710, which is much more reasonable. For daily data, we thus use $\lambda = 0.999$ as the benchmark choice and we do not experiment with values going lower than that. We set the forgetting factor for the model probability updating somewhat lower at $\alpha = 0.99$, which is the recommended value in Koop and Korobilis (2012) and Raftery *et al.* (2010), to allow for a marginally more frequent updating in the model probabilities. Note here also that one could experiment with a few different values, nevertheless, we want to avoid a *'search for the best values'* critique and prefer to use a simple and common parameter setting.

[51] See page 97 in RiskMetrics (1996).

CHF from January 4, 1999 to October 22, 2013. Since we are primarily interested in an out-of-sample forecast evaluation of the combined or model averaged predictions from the two different forecasting agents, we effectively need two *in-sample* or *fitting* periods. That is, we need one *initialisation* period to obtain the agent specific forecasts, $\hat{r}^{\mathsf{F}}_{t+1|t}$ and $\hat{r}^{\mathsf{C}}_{t+1|t}$, and another *calibration* period for the combination weights $\hat{\omega}_t$ in (24) to be determined. We use the first 500 observations (from January 6, 1999 to December 5, 2000) as the *in-sample* period, which in our Kalman Filter setting effectively translates into a *burn-in* period to minimise the influence of the initial values (or priors) on the filtered state vector $\hat{\boldsymbol{\beta}}_{t|t}$.[52] We then use the next 200 observations (from December 6, 2000 to September 11, 2001) to get our estimate $\hat{\omega}_t$ needed to compute our first model averaged out-of-sample forecast for September 12, 2001. We then roll through the rest of the out-of-sample data to update $\hat{\omega}_t$ using a fixed window size of 200 observations and produce (recursively updated) one-step ahead forecasts of the returns. The effective out-of-sample period thus spans from September 12, 2001 to June 30, 2014 (to October 22, 2013 for JPY and CHF), yielding 3339 (respectively 3160) evaluation points.

4.2. Evaluation criteria

We assess the out-of-sample forecast performance of the proposed agent based averaging framework by following the recent literature on forecasting the equity premium. That is, we follow the approach of Rapach *et al.* (2013), Neely *et al.* (2014) and many others and evaluate the forecasts in terms of the Campbell and Thompson (2008) out-of-sample R^2 (denoted by R^2_{os} henceforth) and the Clark and West (2007) Mean Squared Forecast Error (MSFE) adjusted $t-$statistic, which we denote by $\mathrm{CW}-$ statistic. In all our out-of-sample forecast evaluations, we use the forecasts from a driftless random walk (RW) as the benchmark in the statistical tests. To formalise notation, let $\hat{e}^{(\ell)}_{t+1|t}$ denote the (one-step ahead) forecast errors from model ℓ that we

[52] Two observations are lost due to lagging and differencing. Also, note that in our state-space model with given α, λ and κ values there is no in-sample fitting period. So here the first 500 observations are used to minimise the impact of the priors needed to initialise the filter. We actually also estimate the parameters of the full model for each agent type on a rolling and expanding window by OLS. These results were used to obtain a natural benchmark comparison to the Kalman Filter based estimates of the DMA framework (these OLS estimates are available upon request). We judge 500 observations to be large enough to minimise the effect of the priors on the $\hat{\boldsymbol{\beta}}_{t|t}$ estimates.

consider, where $\ell = \{\mathsf{RW}, \mathsf{MC}, \mathsf{F}, \mathsf{C}\}$. These forecast errors are computed as:

$$\hat{e}^{(\ell)}_{t+1|t} = \left(r_{t+1} - \hat{r}^{\ell}_{t+1|t}\right) \tag{31}$$

with corresponding MSFEs being

$$\text{MSFE}_{(\ell)} = \frac{1}{T_{os}} \sum_{t=T_{is}}^{T} \hat{e}^{2(\ell)}_{t+1|t}. \tag{32}$$

The terms T_{os} and T_{is} denote, respectively, the number of out-of-sample and in-sample observations, so that $T_{is} + T_{os} = T$, with T being the full sample size.

The Campbell and Thompson (2008) R^2_{os} is computed as follows. Let $\text{MSFE}_{(\mathsf{MC})}$ be the MSFE from the agent based model combined forecasts and let $\text{MSFE}_{(\text{RW})}$ denote the MSFE from the random walk benchmark model. Then, the R^2_{os} comparing the performance of the MC forecasts to the RW is defined as:

$$R^2_{os} = 1 - \frac{\text{MSFE}_{(\mathsf{MC})}}{\text{MSFE}_{(\text{RW})}}. \tag{33}$$

Intuitively, the R^2_{os} statistic in (33) measures the reduction in the MSFE of the proposed model relative to the benchmark model. When $R^2_{os} > 0$, then this is an indication that the proposed model performs better than the benchmark model in terms of MSFE, while $R^2_{os} < 0$ suggests that the benchmark model performs better.

The Clark and West (2007) MSFE adjusted $t-$statistic is computed as (again assessing the performance of the MC forecasts relative to the RW):

$$\text{CW} - \text{statistic} = -\frac{2}{T_{os}} \sum_{t=T_{is}}^{T} \hat{e}^{(\text{RW})}_{t+1|t} \left(\hat{e}^{(\text{RW})}_{t+1|t} - \hat{e}^{(\mathsf{MC})}_{t+1|t}\right) \tag{34}$$

(see equation 4.1 on page 297 in Clark and West (2007)). Following the suggestion in Clark and West (2007, page 294) , the simplest way to compute the CW − statistic is to form the sequence

$$cw_{t+1} = dm_{t+1} + adj_{t+1} \tag{35}$$

where

$$dm_{t+1} = \hat{e}^{2(\mathsf{RW})}_{t+1|t} - \hat{e}^{2(\mathsf{MC})}_{t+1|t} \tag{36}$$

and

$$adj_{t+1} = \left[\hat{r}^{(\mathsf{RW})}_{t+1|t} - \hat{r}^{(\mathsf{MC})}_{t+1|t}\right]^2. \tag{37}$$

The dm_t term is the standard Diebold and Mariano (1995) sequence that is computed to test for (unconditional) superior predictive ability. The adjustment term adj_t arises due to the nested nature of the models being compared and performs a bias correction (see Clark and West (2007) for more details). The CW − statistic is then computed as

$$\text{CW} - \text{statistic} = \frac{\overline{cw}}{\sqrt{\text{Var}(\overline{cw})}} \tag{38}$$

where $\overline{cw} = T_{os}^{-1} \sum_{t=T_{is}}^{T} cw_{t+1}$ and $\text{Var}(\overline{cw})$ is the variance of the sample mean, which can simply be obtained as the heteroskedasticity and autocorrelation (HAC) robust $t-$statistic on the intercept term from a regression of cw_{t+1} on a constant.[53]

The CW − statistic implements a test of the null hypothesis that the MSFE of the benchmark model is equal to the MSFE of the MC forecasts, against the one sided alternative hypothesis that the benchmark's MSFE is greater than that of the MC. A rejection of the null hypothesis hence suggests that MC forecasts are (on average) *significantly* better than RW forecasts. It should be highlighted here that the CW − statistic is particularly suitable in the given context, as it is designed for a comparison of *nested* (forecasting) models. Our benchmark model is the RW model, which can be obtained from the MC forecasts by restricting $\hat{\boldsymbol{\beta}}^{\mathsf{a}(m)}_{t|t}$ for all $\mathsf{a} = \{\mathsf{F}, \mathsf{C}\}$ in (30) to 0.

In addition to the out-of-sample R^2 of Campbell and Thompson (2008) and the CW − statistic of Clark and West (2007), we also compute the cumulative difference of the MSFEs of the RW and MC forecasts over the out-of-sample period. This cumulative difference (denoted by cumMSFE$_t$) is commonly used in the equity premium forecasting literature as a tool to high-

[53] See also the discussion in Section 2.1 in **?** for more background on this in the context of the traditional Diebold-Mariano (DM) statistic.

light the predictive performance of the model relative to the benchmark over time (see Goyal and Welch (2008) and Rapach *et al.* (2013), among many others). In our setting, this difference is computed as:

$$\text{cumMSFE}_t = \sum_{t=T_{is}}^{T_{os}} \left(\hat{e}_{t+1|t}^{2(\text{RW})} - \hat{e}_{t+1|t}^{2(\text{MC})} \right). \tag{39}$$

A value of cumMSFE_t above zero indicates that the cumulative sum of the squared forecast errors of the RW model are larger than those of the MC forecasts, suggesting that model combined forecasts are more accurate. In general, a rising value in cumMSFE_t means that the MC forecasts produce better predictions than the RW benchmark.

4.3. Forecast evaluation results

In Table 3 we present the one-step ahead out-of-sample forecast evaluation results for the period from September 12, 2001 to June 30, 2014 (to October 22, 2013 for JPY and CHF).[54] The first column in Table 3 shows the models that are fitted, the second column shows the mean squared forecast errors (MSFEs), the third column the MSFEs relative to the RW benchmark, the fourth column the Campbell and Thompson (2008) R^2_{os} (in percent) as defined in (33), and the fifth and sixth columns display the Clark and West (2007) MSFE adjusted $t-$statistic (CW$-$statistic) and its corresponding one-sided $p-$value. Note that the forecasts that are listed here are from the benchmark RW model, the individual chartist and fundamentalist forecasts, as well as the model combined forecasts, where the *'averaging'* weights were obtained from the Granger and Ramanathan (1984) combination regressions in (26).

← Table 3 about here

We can initially notice from the results in Table 3 that chartist forecasts as a whole seem to perform rather poorly when compared to fundamentalist and RW forecasts. This is evident from chartist forecasts producing the largest MSFEs. It is further evident that fundamentalist forecasts generate lower MSFEs than the benchmark RW model, producing out-of-sample R^2 values that are positive and as high as 1.25% for the Swiss Franc, with the Australian Dollar being the only currency with a negative R^2_{os} of -0.07%. From the magnitude of the CW$-$statistics

[54] We group the results for each currency into 4 rows in Table 3.

and their $p-$values, one can see that improvements in forecast accuracy of the fundamentalist forecasts are statistically significant at the 1% level for the Euro, the Yen, the Pound, and the Swiss Franc and significant at the 5% and 10% levels for the Canadian and Australian Dollars, respectively.

Although the above reported out-of-sample R^2 may appear small in magnitude and thus unimportant, we should highlight here that even seemingly low R^2_{os} values can be economically sizeable.[55] In a broader context, one should expect to see only a very small predictive component in exchange rate returns, if foreign exchange markets are believed to be efficient. Also, our R^2_{os} are computed at the daily frequency. To put this magnitude into perspective, Campbell and Thompson (2008) and more recently Neely *et al.* (2014) have shown using monthly data that R^2_{os} values as low as 0.5% (per month) produce economically meaningful predictive results in the sense that *'large'* gains in portfolio performance can be obtained.[56] Overall, we can conclude from the fundamentalist forecast evaluation results that even without optimally combining the forecasts with chartist predictions, statistically significant and sizeable improvements over the random walk benchmark can be obtained.

Looking over the model combined forecast results shown in the last row of each currency grouping in Table 3, we see that, with the exception of the JPY series, all other currency returns benefit from the performance based weighting of the two agent's predictions in the construction of the combined forecasts. The magnitude of this gain can be seen from the higher R^2_{os} values for the EUR, GBP, AUD, CAD and CHF series. For these 5 series, the biggest forecast gain in terms of a higher R^2_{os} is obtained for the Australian Dollar, which increased by approximately 0.36 percentage points and the lowest for the Canadian Dollar, which only improved by about 0.01 percentage points.[57] Note here also that, although chartist forecasts are inferior to funda-

[55] This is a well known result from the equity premium forecasting literature.

[56] What large is here depends on the setting. See the papers by Campbell and Thompson (2008) and Neely *et al.* (2014) for more details on how this is assessed.

[57] It should be stress here again that we *do not* include the observation to be forecasted in the calibration period of the weight ω_t in (24). That is, the combination weight is computed for the first 200 observations, given return forecasts from the two agent types, and the first out-of-sample forecast is then constructed for observation 201. We then roll one observation forward and repeat the fitting and forecasting cycle. This procedure maintains the *'real time'* aspect of the out-of-sample forecast evaluation.

mentalist forecasts when averaged over the full out-of-sample period, there do exist instances where chartist predictions perform better than forecasts based solely on fundamental predictor variables. It is exactly this feature of the MC forecasts that leads to an overall improvement when averaging over the two individual agent based predictions.

To gain a better understanding of the positive (statistical) forecast evaluation results that we obtain, we examine the evolution of the cumulative difference of the MSFEs of the model combined forecasts (relative to the RW benchmark) over time. This cumMSFE_t series, as defined in (39), is plotted in Figure 2 for the 6 currencies of interest. Note here that, because of the shorter out-of-sample evaluation period for the Yen and the Franc, the cumMSFE_t series ends already on October 22, 2013 for these two currencies. Nevertheless, for reasons of comparability with respect to date entries in the figures that we show, we have plotted all series up to June 30, 2014.[58] As a reminder, the cumMSFE_t series is defined such that an increasing value indicates an improvement in the MC predictions relative to the RW benchmark, that is, the RW benchmark produces larger one-step ahead forecast errors.

Looking over the cumMSFE_t series plotted in Figure 2, one can notice the following 4 visually striking features from these plots. First, the cumMSFE_t is (nearly) uniformly above zero for all currencies over the whole out-of-sample evaluation period, except for the AUD series, where it is above zero only up to June 2010, dropping below 0 thereafter. Second, the cumMSFE_t series is (nearly) monotonically increasing for all series up to the September 2008 period, ie., around the time of the Lehman Brothers collapse. Third, the effect of the Lehman Brothers collapse has a *strong* impact on the predictability results for all 6 currencies.[59] From September 2008 until approximately February 2009, the predictability in all 6 series (relative to the RW) experienced a huge boost. This effect is most evident for the Australian Dollar and the Swiss Franc, and least so for the Canadian Dollar. Fourth, since the Lehman Brothers collapse in September 2008, it has become more difficult to outperform the benchmark random walk

[58]Empty (or non-existing) values of the shorter out-of-sample values for the JPY and CHF series are simply set to "NAN" to plot the series. We also set a common $y-$axis scale over the interval $[-5, 35]$ to facilitate the comparison of the magnitudes across the 6 currencies.

[59]A similar result is found in Buncic and Moretto (2014) in a forecast evaluation using LME copper data.

predictions. This conclusion can be reached from the overall downward *'trend'* that is visible in the cumMSFE$_t$ series from approximately February 2009 until the end of the sample period in June 2014 (October 2013).

The single most peculiar visual result of the cumMSFE$_t$ series is for the Australian Dollar, where the forecast accuracy of the model combined predictions initially improves dramatically, with the cumMSFE$_t$ increasing from a value of approximately 19 on October 9, 2008 to nearly 28 on October 10, 2008, then dropping back to 19 on October 13, and decreasing rapidly thereafter. By October 21, 2008, the cumMSFE$_t$ had dropped below 6. For the Euro and the Canadian Dollar, similar but less accentuated swings in the cumMSFE$_t$ are visible. For the Euro, the cumMSFE$_t$ increased from a value of around 12 on September 9, 2008 to nearly 20 by October 6, 2008, dropping back to just below 13 on November 21, 2008, and rising up to 21 by December 19, 2008, before gradually dropping, reaching a value of 15 by May 2009. For the Canadian dollar, although at an overall smaller cumMSFE$_t$ magnitude, also fairly abrupt movements are visible. For instance, the cumMSFE$_t$ stood at about 5 on September 29, 2008, and then doubled to over 10 by October 10, 2008. By February 2010, it had declined to a value of 2. The Yen, Pound and Franc experienced more stable and seemingly permanent increases in predictability over the September 2008 to May 2009 period. This is highlighted by the cumMSFE$_t$ series increasing from values of around 6 to 14, 5 to 12, and 15 to 25, for the JPY, GBP and CHF, respectively over this time span. All in all, it should be clear from the visual analysis that we presented here that around the time of the Lehman Brothers collapse a substantial boost in exchange rate predictability was realised.[60]

To learn about the influence of chartist agents, we show the time series evolution of the weight function $\hat{\omega}_t$, as defined in (24), in Figure 3. $\hat{\omega}_t$ is plotted as a red solid line in Figure 3. We also create an indicator variable that is equal to 1 if $\hat{\omega}_t > 0.5$. This indicator variable is drawn as a gray shaded area in Figure 3. The time series evolution of the weight functions

[60]It should be stressed here that several papers find some predictability of exchange rates before the financial crisis, with their results, nevertheless, breaking down *after* 2008/2009 (see, for instance,Adrian *et al.* (2010), Molodtsova *et al.* (2011), and Molodtsova and Papell (2012), and references therein). Our improved predictability result over the crisis period is thus entirely new. It should be stressed here, nevertheless, that these studies use standard low frequency data, thus face an entirely different conditioning set upon which the forecasts are based.

portray a number of interesting insights. First, we can notice that the weight of chartists agents in the model combined forecasts varies considerably between the 6 currencies of interest. The influence of chartist agents was rather low for the first four years into the out-of-sample period for the EUR, the AUD and the CHF, with some episodes of importance for GBP, CAD and JPY over the 2002 to 2003 period. However, from the mid to end of 2009 onwards, $\hat{\omega}_t$ began to increase steadily. Recall that this period coincides with the strong rebound in global equity prices that followed the bottom of the bear market in March 2009 and also the flow on effects of the implementation of quantitative easing in late November 2008 in the US. This increase in chartist weights is particularly noticeable for the EUR, the GBP and AUD series. For the Yen, chartist agents did not bear any important effect on the forecasts until the beginning of March 2011, while for the Canadian Dollar, the period following the bottom of the equity bear market in March 2009 to March 2011 was, apart from a short time interval from March 2010 to May 2010, largely driven by chartist forecasts.

What is interesting to see from the model combination weights is that around the period of the Lehman Brothers collapse, where the Pound as well as the Australian and Canadian Dollars depreciated fairly rapidly, and where one would thus expect to see fairly strong momentum in the currencies, the weight given to chartist forecasts was close to zero for the Pound and the Canadian Dollar. For these two currencies, the information contained in the fundamental predictors superseded the momentum ones used by chartist agents. For the Australian Dollar, the situation was somewhat different in the sense that the influence of chartist predictions in the MC forecasts already started to increase in July 2008, reaching a weight of 1 by mid September 2008, then episodically dropping to zero from October 9 to October 13, 2008, before jumping back up to 1 by October 17, 2008, and remaining more or less at that level until May 2013. For the Australian Dollar, the performance of the overall model combined forecast was actually worse than the RW benchmark for the entire post February 2009 period.[61] Nevertheless, it is clear that the *'good'* performance of the fundamentalist agent based forecasts had diminished

[61]One needs to be careful not to interpret too much into the meaning of the weights, as it is clear that both fundamentalist as well as chartist did rather poorly compared to the benchmark random walk forecast over this time period.

substantially since February 2009.[62]

4.4. Forecasting performance at longer horizons

Given the overall positive results at the one day ahead horizon, we now turn to assess the forecasting performance of our proposed fundamentalist and chartist model combined predictions at horizons of 2 up to 5 days ahead. To construct multiple-step ahead out-of-sample forecasts from the agent based model combination, we implement the so-called '*direct*' forecasting approach.[63] That is, we re-formulate the relation in (24) for the general $h-$step ahead relation as:

$$\hat{r}^{\mathsf{MC}}_{t+h|t} = (1-\hat{\omega}_{t,h})\hat{r}^{\mathsf{F}}_{t+h|t} + \hat{\omega}_{t,h}\hat{r}^{\mathsf{C}}_{t+h|t} \tag{40}$$

where $\hat{r}^{\mathsf{F}}_{t+h|t}$ and $\hat{r}^{\mathsf{C}}_{t+h|t}$ are respectively, the individual $h-$step ahead ($h-$period holding return) fundamentalist and chartist forecasts, and the ($h-$step) ahead weighting function $\hat{\omega}_{t,h}$ is obtained, analogous to (26), from the regression:

$$(r_{t-h:t} - \hat{r}^{\mathsf{F}}_{t|t-h}) = \omega_{t,h}(\hat{r}^{\mathsf{C}}_{t|t-h} - \hat{r}^{\mathsf{F}}_{t|t-h}) + \nu_t \tag{41}$$

with $r_{t-h:t} = 100(P_t/P_{t-h} - 1)$ denoting the $h-$period holding return of the currency of interest.

4.4.1. Computing agent specific multiple period ahead out-of-sample forecasts

We also use the direct forecasting approach to construct individual, agent specific $h-$period ahead forecasts from the DMA framework. To do this, we re-write the relation in (3) (again

[62]At this point, it is not clear why this was the case. One could argue that the exceptional performance over the first half of the out-of-sample period was driven by carry trades. Nevertheless, one could make a counter argument that a carry trade strategy would have been equally as successful after the Lehman Brothers collapse, as the interest differentials remained, if not widened at that time. One should keep in mind though, that a key ingredient to carry trades is a certain level of risk appetite which had clearly diminished after the Lehman Brothers. Another key point that needs to be looked at is whether the fixed α and λ parameter setting that we maintain should be relaxed over this period. This is computationally demanding and could be investigated with a smaller set of predictor variables. We leave these issues aside for now, and plan to address them in further research.

[63]See Clements and Hendry (1996), Chevillon and Hendry (2005), Marcellino *et al.* (2006), Chevillon (2007), and Pesaran *et al.* (2011), among others, for a motivation, evaluation and comparison of the direct forecasting approach to iterated forecasts.

using the general y_t and $\mathbf{x}_t$ notation as in Section 2.1) as

$$y_t = \mathbf{x}_{t-h}^{(m)} \boldsymbol{\beta}_{t,h}^{(m)} + u_t^{(m)} \tag{42a}$$

$$\boldsymbol{\beta}_{t,h}^{(m)} = \boldsymbol{\beta}_{t-1,h}^{(m)} + \boldsymbol{\epsilon}_t^{(m)}, \tag{42b}$$

where the h subscript in $\boldsymbol{\beta}_{t,h}^{(m)}$ signifies the relation to the $h-$period lagged value of $\mathbf{x}_t$. Using the same Kalman Filter recursions as in (5), but now on the $h-$period lagged relation as specified in (42) yields filtered estimates of the latent states (for each model and agent type $\mathsf{a} = \{\mathsf{F}, \mathsf{C}\}$), that is, $\hat{\boldsymbol{\beta}}_{t|t,h}^{\mathsf{a}(m)}$.

Given $\hat{\boldsymbol{\beta}}_{t|t,h}^{\mathsf{a}(m)}$, the DMA based agent specific $h-$step ahead forecasts are then computed as:

$$\hat{r}_{t+h|t}^{\mathsf{a}} = \sum_{m=1}^{M} \mathbf{x}_t^{\mathsf{a}(m)} \hat{\boldsymbol{\beta}}_{t|t,h}^{\mathsf{a}(m)} \pi_{t+h|t}^{\mathsf{a}(m)} \tag{43a}$$

for all $t = T_{is}, \ldots, T$, where, again due to the random walk evolution of the latent state vector $\boldsymbol{\beta}_{t+h,h}^{\mathsf{a}(m)}$, the best forecast of $\boldsymbol{\beta}_{t+h,h}^{\mathsf{a}(m)}$ is its last observed filtered estimate, that is, $\mathbb{E}_t(\boldsymbol{\beta}_{t+h,h}^{\mathsf{a}(m)}) = \hat{\boldsymbol{\beta}}_{t|t,h}^{\mathsf{a}(m)}$. The $h-$step ahead agent specific predictive model probabilities at time t are computed from

$$\pi_{t+h|t}^{\mathsf{a}(m)} = \frac{\pi_{t|t}^{\mathsf{a}\alpha(m)}}{\sum_{j=1}^{M} \pi_{t|t}^{\mathsf{a}\alpha(j)}}, \tag{44}$$

with $\pi_{t|t}^{\mathsf{a}(m)}$ being the (agent specific) filtered model probability, analogous to the definition given in (10b), $\forall \mathsf{a} = \{\mathsf{F}, \mathsf{C}\}$.

4.4.2. Multiple-step ahead forecast evaluation

We use the same calibration for the λ, α and κ parameters that were used in the one-step ahead prediction setting to implement the Kalman Filter recursions.[64] The results for forecast horizons $h = 2, 3, 4$, and 5 days ahead are reported in Table 4. To avoid clutter in the table, we only

[64]We simply leave these parameters fixed again to avoid concerns related to *'fishing for the best out-of-sample results'*. One could again try to optimise over these parameters, but due to the computational burden, we do not consider this here.

report results for the model combined forecasts for each exchange rate series at the four different forecast horizons.[65] Also, since $h-$step ahead forecast errors will be MA($h-1$) processes in general (ie., will be moving averages and therefore autocorrelated of order $h-1$) which affects the CW$-$statistic, we use a HAC robust variance for $\text{Var}(\overline{cw})$ in (38). More specifically, we follow the recommendation of Andrews and Monahan (1992) and employ a data driven bandwidth using a Quadratic Spectral (QS) Kernel with a *'pre-whitening'* step, where we choose the (optimal) bandwidth parameter with an AR(1) as the approximating model (see equation 3.5 in Andrews and Monahan (1992)).[66] The last four columns in Table 4 are the same as the last four columns in Table 3. Columns 2 and 3 in Table 4 show respectively the MSFEs of the RW benchmark and the MC predictions. The first column lists the various forecast horizons.

← Table 4 about here

From the results reported in Table 4 it is evident that the forecast performance of the MC predictions — relative to the random walk forecast — remains in tact for forecast horizons up to 2 days ahead for all currencies (except for the Australian Dollar) generating out-of-sample R^2 values of 0.27, 0.18, 0.12, 0.16 and 0.18 percent for the EUR, JPY, GBP, CAD and CHF exchange rate series, respectively. Moreover, these improvements are statistically significant at the 5% level for the EUR, JPY, GBP, and CHF, and at the 10% level for the CAD. What is interesting to see is that for a holding period return of even 3 days, improvements in the R^2_{os} are still realised for the EUR ($R^2_{os} = 0.30\%$), the CAD ($R^2_{os} = 0.33\%$) and also for the AUD series, albeit with a lower R^2_{os} of 0.08, and are overall significant at the 10% level. For the EUR, CAD and the AUD series, forecasting the three day holding return $r_{t:t+3}$ yields in fact higher R^2_{os} values than forecasts two days ahead. At forecast horizons of 4 and 5 days ahead, the MC predictions start to produce consistently worse forecasts than the benchmark random walk model across all 6 exchange rates, as the forecast horizon increases.

[65] At horizons 2 to 5, we also find that for the majority of exchange rates, the combined forecasts improve on the fundamental ones. The exception is again the JPY series, at all forecast horizons and also the CAD series for forecast horizons above 2. Further details, if needed, are available upon request from the authors.

[66] That is, to pre-whiten the cw_{t+1} series, we first fit an ARMA(1, 1) to cw_{t+1} and then use the QS Kernel with the bandwidth parameter set to $1.3221\left(\hat{\alpha}(2)T_{os}\right)^{1/5}$, where $\hat{\alpha}(2) = 4\hat{\rho}^2/(1-\hat{\rho})^4$ and $\hat{\rho}$ is the AR(1) parameter estimate obtained from an AR(1) regression of the (pre-whitened) residual series obtained from the ARMA(1, 1) model fitted to cw_{t+1}. To obtained the HAC variance, we then *'re-colour'* again with the ratio of the square of the ARMA lag polynomials (see Andrews and Monahan, 1992 for more details on the exact computations).

As was done for the one-step ahead forecasts, we again show plots of the cumMSFE$_t$ of the model combined forecasts, for $h = 2, \ldots, 5$ in Figure 4 to visualise how the performance of the forecasts evolves over the out-of-sample evaluation period relative to the RW benchmark. In order to facilitate the comparison across the various forecast horizons, we plot all cumMSFE$_t$ for each exchange rate and considered forecast horizon in one subfigure in Figure 4 and off-set the various horizons so that they can be plotted in the same graph.[67] From the time series plots of the cumMSFE$_t$ series in Figure 4, a number of interesting features stand out. First, the increase in the cumMSFE$_t$ around (and following) the time of the Lehman Brothers collapse remains visible for forecast horizons of up to 5 days ahead for the EUR, GBP, CAD series, and also, but to a lesser extent, for the CHF series. For the Australian Dollar, a substantial worsening in the predictive ability with respect to the RW benchmark can be seen, most evidently for forecast horizons 5, 4 and 2 days ahead. What is perhaps somewhat surprising to see from the cumMSFE$_t$ plot for the Australian Dollar is the improved forecast result at the 3 days ahead horizon. This is evident from the (red) cumMSFE$_t$ corresponding to $h = 3$ being consistently above the 0 line in Figure 4. The Lehman Brothers collapse seems to have had a rather positive impact on the predictive performance for 3 day ahead forecasts, and is therefore more inline with the effects experienced by the other 5 currencies.

← Figure 4 about here

Second, the improved forecast performance at the 3 steps-ahead horizon which were found from the statistical evaluation results in Table 4 for the EUR and CAD are largely driven by the strong performance of the MC forecasts around the time of the Lehman Brothers collapse. This can be seen from the persistent upward jump in the cumMSFE$_t$ around the September 2008 period. What is interesting to highlight here is that the forecast performance of the MC predictions remained fairly stable after the Lehman Brothers collapse, which was not the case for the Canadian Dollar at the one-step ahead horizon, while for the EUR the cumMSFE$_t$ series decreased somewhat, indicative of a mild worsening in forecast performance with respect to the RW benchmark. Third, for the Swiss Franc, the MC predictions are consistently superior to

[67] That is, we again pad the first few entries of the $h = 3, 4$ and 5 step-ahead horizons with "NANs" so that the dimensions of the forecast vectors are the same in the date $x-$axis dimension.

RW forecasts over the period from September 2008 to September 2011 for all 5 forecast horizons that we consider. Moreover, a noticeable build-up in the predictive improvement seems to occur from approximately November 2010 until about August 2011. This time period coincides with the strong appreciation period that we identified for the Swiss Franc from Figure 1. The momentum in the currency over this time period created a substantial opportunity to outperform the no-change random walk prediction, even at forecast horizons of up to 5 days ahead. This is an interesting result that has not been reported in the exchange rate forecasting literature thus far.

In summary of the multiple-step ahead forecast evaluation, we can conclude that the predictability results that we obtained at the one-step horizon carry over to forecasts of up to 3 days ahead for most of the currencies that we analyse. A visual analysis of the time evolution of the cumMSFE$_t$ series highlights that a substantial part in the predictive performance gain of the MC forecasts is realised during the September 2008 to February 2009 period, ie., around the time of the Lehman Brothers collapse and a few months after. For the predicability results of the Swiss Franc, particulary at forecast horizons of more than one day ahead, this key time period lasts until about August 2011, that is, approximately from the Lehman Brothers collapse until the time speculation of the imposition of a cap on the EUR/CHF pair first started to circulate.

5. Conclusion

We employ an empirical heterogeneous agent model consisting of fundamentalists and chartist agents to forecast 6 of the most frequently traded currencies using daily data over the sample period from January 1999 to June 2014. More specifically, we use a (time varying) model combination approach to optimally average the forecasts from individual fundamentalist and chartist agents, where individual fundamentalist and chartist predictions are constructed using the recently proposed flexible DMA framework. The 6 currencies that are used in the forecast evaluation are the Euro, the Yen, the British Pound, the Australian and Canadian Dollars, and the Swiss Franc. Given the daily nature of our forecast construction, we use financial data as proxy variables for the macroeconomic predictors used by fundamentalist agents. These fundamental predictors contain level, slope and curvature yield curve factors, equity indices as well as data related to global trade and risk aversion. To model the behaviour of chartist agents, we construct chartist predictor variables consisting of moving averages and momentum to capture the well known trend following behaviour of chartists and also relative strength indices to capture overbought or oversold conditions in the foreign exchange market and reversal trading strategies.

Covering an out-of-sample period from September 2001 to June 2014 we show that forecasts from our empirical heterogeneous agent model combined predictions significantly outperform the forecasts from the standard random walk benchmark model for all 6 currencies that are considered. The (daily) Campbell and Thompson (2008) out-of-sample R^2 can be as high as 1.41%, 1.07%, 0.99%, and 0.74% for the Franc, the Euro, the Pound and the Yen series, and is somewhat lower for the Australian and Canadian Dollars at 0.29% and 0.24%. Statistical tests of significance show that these are significant at the 10% level for the Australian and Canadian Dollars, and even at the 1% level for the remaining 4 currencies. Forecast gains relative to the random walk benchmark remain statistically significant at the 10% level for forecasts up to three days ahead for the CHF, CAD, AUD, and EUR series, producing out-of-sample R^2 values as high as 0.33% for the Canadian Dollar. It should be stressed here again that these out-of-sample R^2 values are computed at the daily frequency, keeping in mind the results of Campbell and Thompson (2008) that even a small out-of-sample R^2 of 0.5% at a monthly frequency generates economically meaningful returns to investors.

From the time series evolution of the cumulative difference of the MSFEs of the RW and model combined forecasts (abbreviated as cumMSFE_t) it is evident that, at the one-step ahead horizon, the cumMSFE_t remains above zero for the entire out-of-sample period for all currencies, with the only exception being the Australian Dollar, where it drops below 0 in June 2010, remaining below 0 until the end of the evaluation period in June 2014. What this highlights is the consistency of the improvement of the model combined forecasts over the entire sample period and for (nearly) all currencies that we consider. From the cumMSFE$_t$ plots it is further evident that the period around and following the Lehman Brothers collapse in September 2008 until about February 2009 was marked by a drastic improvement in the agent based MC forecasts relative to the RW benchmark. For the Australian Dollar, this period is marked by some erratic behaviour in the cumMSFE$_t$ showing initially a jump upwards in the series, indicative of a substantial improvement in the predictive ability relative to the RW model, but followed by a sharp drop in December 2008, and decreasing further thereafter. This emphasises the changing nature of the forecast environment and warrants further investigation into the mechanisms that lead to this behaviour. What is interesting to see from the cumMSFE$_t$ plots at the various forecast horizons is that the influence of the Lehman Brothers collapse remains strongly visible even at horizons of up to 5 days ahead.

Finally, the evolution of the chartist weight function (denoted by $\hat{\omega}_t$) shows that for large parts of the first half of the out-of-sample evaluation period, mainly fundamentalist agents (or predictor variables) were contributing to the improvement in the MC forecasts over the RW model. Nevertheless, for the last third of the out-of-sample period, the influence of chartist forecasts in the MC predictions increased noticeably, particularly for the GBP, EUR and AUD exchange rate series, and to a lesser extent for the CHF and JPY. Over this period, these currencies experienced a fairly strong appreciation relative to the US Dollar, generating a considerable amount of momentum in returns MC.

References

Adrian, Tobias, Erkko Etula and Hyun Song Shin (2010): "Risk Appetite and Exchange Rates," *Federal Reserve Bank of New York Staff Reports No. 361*, Federal Reserve Bank of New York.

Allen, Helen and Mark P. Taylor (1990): "Charts, Noise and Fundamentals in the Foreign Exchange Market," *The Economic Journal*, **100**(1), 4959.

Altavilla, C. and P. D. Grauwe (2010): "Forecasting and combining models of exchange rate determination," *Applied Economics*, **42**, 3455–3480.

Andrews, Donald W. K. and J. Christopher Monahan (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, **60**(4), 953–966.

Aramov, D. (2002): "Stock return predictability and model uncertainty," *Journal of Financial Economics*, **64**, 423–458.

Bates, John M. and Clive W. J. Granger (1969): "The Combination of Forecasts," *Operations Research Quarterly*, **20**(4), 451–468.

Baumeister, Christiane and Lutz Kilian (2014): "What Central Bankers Need to Know about Forecasting Oil Prices," *International Economic Review*, **55**(3), 869–889.

Bekaert, Geert and Robert J. Hodrick (1992): "Characterizing predictable components in excess returns on equity and foreign exchange markets." *Journal of Finance*, **47**(2), 467–509.

Bhar, Ramprasad and Girijasankar Mallik (2013): "Inflation uncertainty, growth uncertainty, oil prices, and output growth in the UK," *Empirical Economics*, **45**(3), 1333–1350.

Bilson, J. F. O. (1990): *Technical currency trading*, The currency-hedging debate, London, International Financing Review Publishing.

BIS (2013): *"Foreign exchange turnover in April 2013: Preliminary global results,"* Triennial Central Bank Survey, Bank for International Settlements. Available from: http://www.bis.org/publ/rpfx13fx.pdf.

Bodie, Z., A. Kane and A.J. Marcus (2002): *Investments*, McGraw-Hill/Irwin, Boston.

Brand, Claus, Daniel Buncic and Jarkko Turunen (2010): "The Impact of ECB Monetary Policy Decisions and Communication on the Yield Curve," *Journal of the European Economic Association*, **8**(6), 1266–1298.

Brock, William A. and Cars H. Hommes (1997): "A Rational Route to Randomness," *Econometrica*, **65**(5), 1059–1095.

——— (1998): "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model," *Journal of Economic Dynamics and Control*, **22**(8-9), 1235–1274.

Brunnermeier, Markus K., Stefan Nagel and Lasse H. Pedersen (2009): "Carry Trades and Currency Crashes," in *NBER Macroeconomics Annual 2008*, edited by Daron Acemoglu, Kenneth Rogoff and Michael Woodford, University of Chicago Press, Volume 23, 313–347s.

Buncic, Daniel (2012): "Understanding forecast failure of ESTAR models of real exchange rates," *Empirical Economics*, **34**(1), 399–426.

Buncic, Daniel and Martin Melecky (2014): "Equilibrium credit: The reference point for macroprudential supervisors," *Journal of Banking and Finance*, **41**(April), 135–154.

Buncic, Daniel and Carlo Moretto (2014): "Forecasting Copper Prices with Dynamic Averaging and Selection Models," *Discussion Paper No. 2014-30*, School of Economics and Political Sciences, University of St. Gallen.

Campbell, John Y. and Samuel B. Thompson (2008): "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies*, **21**(4), 1509–1531.

Cheung, Y. W., M. D. Chinn and A. G. Pascual (2005): *What do we know about recent exchange rate models? In-sample fit and out-of-sample performance evaluated.*, Exchange Rate Modelling: Where do we stand? MIT Press for CESifo, Cambridge, MA, 239–276.

Chevillon, Guillaume (2007): "Direct multi-step Estimation and Forecasting," *Journal of Economic Surveys*, **21**(4), 746–785.

Chevillon, Guillaume and David F. Hendry (2005): "Non-parametric direct multi-step estimation for forecasting economic processes," *International Journal of Forecasting*, **21**(2), 201–218.

Chong, Terence Tai-Leung and Wing-Kam Ng (2008): "Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30," *Applied Economics Letters*, **15**(14), 1111–1114.

Clarida, Richard H., Lucio Sarno, Mark P. Taylor and Giorgio Valente (2003): "The out-of-Sample Success of Term Structure Models as Exchange Rate Predictors: A Step Beyond," *Journal of International Economics*, **60**(1), 61–83.

Clark, Todd E. and Kenneth D. West (2007): "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics*, **138**(1), 291–311.

Clements, Michael P. and David F. Hendry (1996): "Multi-step Estimation for Forecasting," *Oxford Bulletin of Economics and Statistics*, **58**(4), 657–84.

Collin-Dufresne, Pierre, Robert S. Goldstein and J. Spencer Martin (2001): "The Determinants of Credit Spread Changes," *The Journal of Finance*, **56**(6), 2177–2207.

Cremers, K. J. Martijn (2002): "Stock Return Predictability: A Bayesian Model Selection Perspective," *The Review of Financial Studies*, **15**(4), 1223–1249.

Croushore, Dean and Tom Stark (2001): "A funny thing happened on the way to the data bank: A real time data set for macroeconomists," *Federal Reserve Bank of Philadelphia - Business Review*, (September/October), 15–27.

De Grauwe, Paul and Marianna Grimaldi (2005): "Heterogeneity of Agents, Transaction Costs and the Exchange Rate," *Journal of Economic Dynamics and Control*, **29**(4), 691–719.

——— (2006): "Exchange Rate Puzzles: A Tale of Switching Attractors," *European Economic Review*, **50**(1), 1–33.

Diebold, Francis X. and Roberto S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, **13**(1), 253–263.

Diebold, Francis X., Glenn D. Rudebusch and S. Boragan Aruoba (2006): "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," *Journal of Econometrics*, **131**(1-2), 309–338.

Engel, Charles M. and James D. Hamilton (1990): "Long Swings in the Dollar: Are They in the Data and Do Markets Know It," *American Economic Review*, **80**(4), 689–713.

Gartley, H. M. (1935): *Profits in the Stock Market*, Lambert-Gann Publishing, Pomerory, Wash.

Goyal, Amit and Ivo Welch (2008): "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, **21**(4), 1455–1508.

Granger, Clive W. J. and Ramu Ramanathan (1984): "Improved methods of combining forecasts," *Journal of Forecasting*, **3**(2), 197–204.

Gürkaynak, Refet S., Andrew T. Levin and Eric T. Swanson (2005): "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, **1**(1), 55–93.

Gürkaynak, Refet S., Brian Sack and Jonathan H. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, **54**(8), 2291–2304.

Hamilton, James D. (2009): "Causes and Consequences of the Oil Shock of 2007-08," *Brookings Papers on Economic Activity*, **40**(1), 215–283.

Hansen, Bruce E. (2008): "Least-squares forecast averaging," *Journal of Econometrics*, **146**(2), 342–350.

Harvey, A. C. (1989): "Forecasts of Economic Growth from the Bond and Stock Markets," *Financial Analysts Journal*, **45, No. 5**, 38–45.

Harvey, C. R. (1993): "Term Structure Forecasts Economic Growth," *Financial Analysts Journal*, **49, 3**, 6–8.

He, Yanan, Shouyang Wang and Kin Keung Lai (2010): "Global economic activity and crude oil prices: A cointegration analysis," *Energy Economics*, **32**(4), 868–876.

Kilian, Lutz (2008): "The Economic Effects of Energy Price Shocks," *Journal of Economic Literature*, **46**(4), 871–909.

Klaassen, Franc (2005): "Long swings in exchange rates: Are they really in the data?" *Journal of Business and Economic Statistics*, **23**(1), 87–95.

Koop, Gary and Dimitris Korobilis (2012): "Forecasting Inflation Using Dynamic Model Averaging," *International Economic Review*, **53**(3), 867–886.

Koop, G. and S. Potter (2004): "Forecasting in dynamic factor models using Bayesian model averaging," *Econometrics Journal*, **7**, 550–565.

Lane, Philip R. (2012): "The European Sovereign Debt Crisis," *Journal of Economic Perspectives*, **26**(3), 49–68.

Lardic, Sandrine and Valerie Mignon (2008): "Oil prices and economic activity: An asymmetric cointegration approach," *Energy Economics*, **30**(3), 847–855.

Lui, Yu-Hon and David Mole (1999): "The Use of Fundamental and Technical Analysis by Foreign Exchange Dealers: Hong Kong Evidence," *Journal of International Money and Finance*, **17**(3), 535–545.

Malkhozov, Aytek, Philippe Mueller, Andrea Vedolin and Gyuri Venter (2014): "Funding Liquidity CAPM: International Evidence," *SSRN Working Paper No. 2420772*. Available from: http://ssrn.com/abstract=2420772.

Marcellino, Massimiliano, James H. Stock and Mark W. Watson (2006): "A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series," *Journal of Econometrics*, **135**(1-2), 499–526.

Mark, Nelson C. (1995): "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *The American Economic Review*, **85**(1), 201–218.

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

Menkhoff, Lukas (1998): "The Noise Trading Approach Questionnaire Evidence from Foreign Exchange," *Journal of International Money and Finance*, **17**(3), 547–564.

Molodtsova, Tanya., Alex. Nikolsko-Rzhevskyy and David H. Papell (2011): "Taylor Rules and the Euro." *Journal of Money, Credit and Banking*, **43**(2-3), 535–552.

Molodtsova, Tanya and David H. Papell (2012): "Taylor Rule Exchange Rate Forecasting During the Financial Crisis," *NBER Working Paper No. 18330*, National Bureau of Economic Research.

Neely, Christopher J., David E. Rapach, Jun Tu and Guofu Zhou (2014): "Forecasting the Equity Risk Premium: The Role of Technical Indicators," *Management Science*, **60**(7), 1772–1791.

Pan, Jun and Kenneth J. Singleton (2008): "Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads," *The Journal of Finance*, **63**(5), 2345–2384.

Pesaran, M. Hashem, Andreas Pick and Allan Timmermann (2011): "Variable selection, estimation and inference for multi-period forecasting problems," *Journal of Econometrics*, **164**(1), 173–187.

Raftery, Adrian E., Miroslav Kárný and Pavel Ettler (2010): "Online Prediction Under Model Uncertainty via Dynamic Model Averaging: Application to a Cold Rolling Mill," *Technometrics*, **52**(1), 52–66.

Rapach, David E., Jack K. Strauss and Guofu Zhou (2013): "International Stock Return Predictability: What Is the Role of the United States?" *The Journal of Finance*, **68**(4), 1633–1662.

RiskMetrics (1996): *"RiskMetrics: Technical Document, 4th Edition,"* J.P.Morgan/Reuters. Available from: http://www.riskmetrics.com/system/files/private/td4e.pdf.

Rodríguez-González, Alejandro, Fernando Guldrís-Iglesias, Ricardo Colomo-Palacios *et al.* (2010): "Improving Trading Systems Using the RSI Financial Indicator and Neural Networks," in *Knowledge Management and Acquisition for Smart Systems and Services*, edited by Byeong-Ho Kang and Debbie Richards, Springer-Varlag, Berlin Heidelberg, 27–37.

Rogoff, Kenneth S. (1996): "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, **34**(2), 647–668.

Rossi, Barbara (2013): “Exchange Rate Predictability,” *Journal of Economic Literature*, **51**(4), 1063–1119.

Rudebusch, Glenn D. and Tao Wu (2008): “A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy,” *The Economic Journal*, **118**(530), 906–926.

Sarno, Lucio and Mark Taylor (2003): *The economics of exchange rates*, Cambridge.

Stock, J. H. and M. W. Watson (1998): “Median Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model,” *Journal of the American Statistical Association*, **93**(441), 349–358.

Sullivan, Ryan, Allan Timmermann and Halbert White (1999): “Data-Snooping, Technical Trading Rule Performance, and the Bootstrap,” *The Journal of Finance*, **54**(5), 1647–1691.

Svensson, Lars E. O. (1994): “Estimating and Interpreting Forward Rates: Sweden 1992-4,” *NBER Working Paper No. 4871*, National Bureau of Economic Research.

Taylor, Mark P. and Helen Allen (1992): “The Use of Technical Analysis in the Foreign Exchange Market,” *Journal of International Money and Finance*, **11**(3), 304–314.

Wilder, J. W. (1978): *New Concepts in Technical Trading Systems*, Trend Research.

Wright, Jonathan H. (2008): “Bayesian Model Averaging and exchange rate forecasts,” *Journal of Econometrics*, **146**(2), 329–341.

Figures and Tables

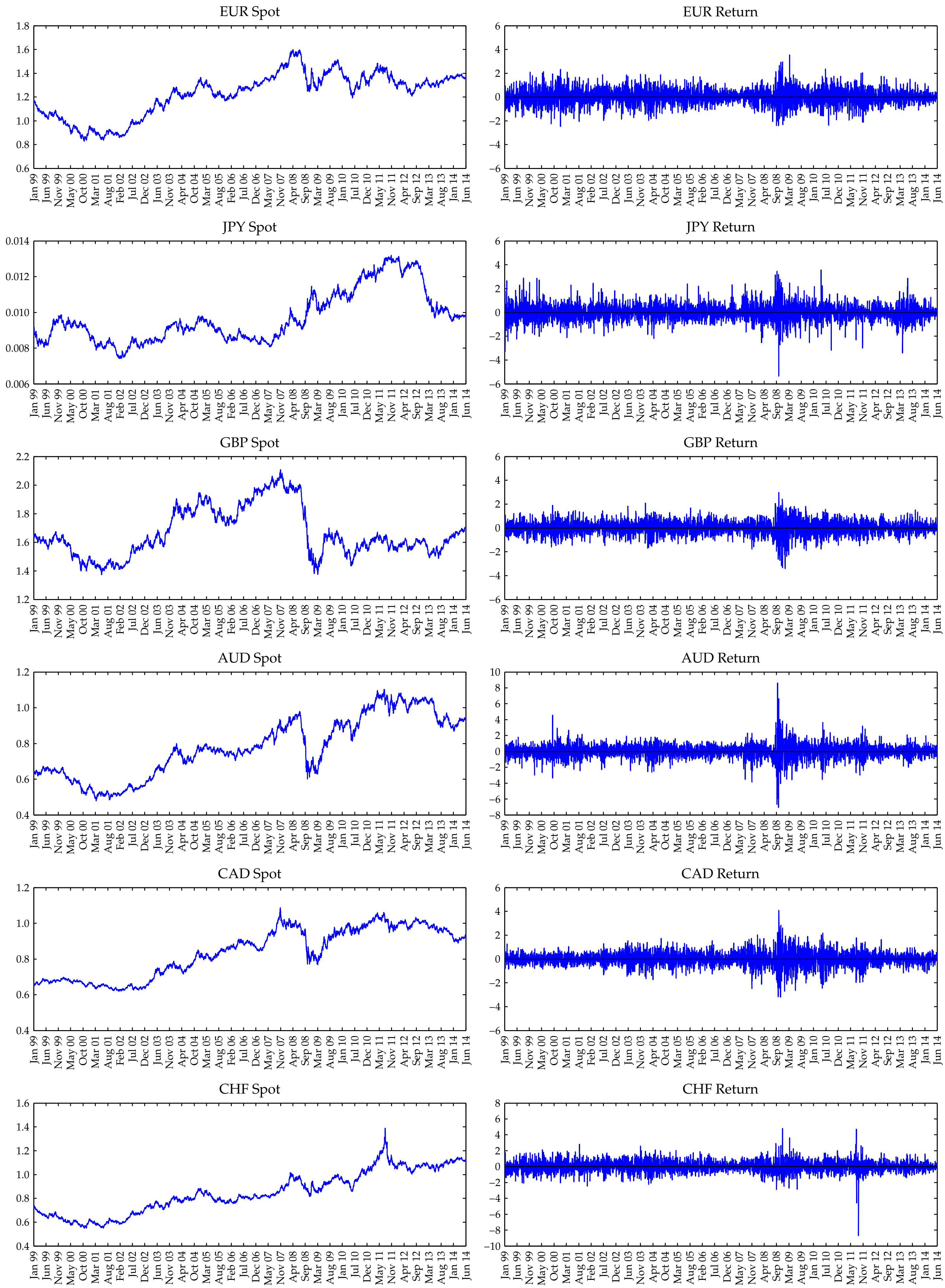

Figure 1: Time series plots of the six exchange rates of interest (full sample period from January 4, 1999 to June 30, 2014). Left Panel shows the (raw) level series. Right Panel shows the return series used in the forecast evaluation.

Table 1: Summary statistics of all exchange rate returns and fundamental predictor variables

Variable Name	Description and Data source	Mean	Median	Std.Dev	Skew	Kurt	Min	Max	ACF(1)
r_t^{EUR}	Exchange rate returns: $r_t^i = 100(P_t^i/P_{t-1}^i - 1),\ \forall i = \{$EUR, JPY, GBP, AUD, CAD, CHF$\}$. All exchange rate data are taken from Bloomberg, where Prices P_t^i are PX_Last entries for the i^{th} currency, using the New York 5pm snap time stamp.	0.0056	0.0130	0.6359	0.0683	4.3251	−2.4870	3.5445	−0.0158
r_t^{JPY}		0.0046	0.0082	0.6560	0.0316	6.4730	−5.3555	3.5659	−0.0411
r_t^{GBP}		0.0023	0.0143	0.5591	−0.2367	5.2895	−3.4123	2.9633	0.0272
r_t^{AUD}		0.0138	0.0507	0.8355	−0.2062	12.6416	−7.0500	8.6215	−0.0431
r_t^{CAD}		0.0105	0.0129	0.5667	−0.1146	6.3163	−3.2017	4.0791	−0.0461
r_t^{CHF}		0.0131	0.0088	0.6932	−0.3031	11.3561	−8.6881	4.8037	−0.0380
$\Delta(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^{\mathrm{EU}})$	Differences in Level, Slope and Curvature factors, relative to the US. Factors are constructed as described in (13). US yield curve data are taken from the Gürkaynak *et al.* (2007) database. For the Euro area, yields are combined from Bundesbank and ECB yield curve data, using the Nelson-Siegel-Svensson parameter estimates. Yield curve data for CH and JP are constructed using the Nelson-Siegel-Svensson parameter estimates from Malkhozov *et al.* (2014). All yield curve data (except for CH and JP) were taken from the respective central bank websites at: Bundesbank: http://www.bundesbank.de/Navigation/EN/Statistics/Time_series_databases/Macro_economic_time_series/its_list_node.html?listId=www_s140_it03c ECB: https://www.ecb.europa.eu/stats/money/yc/data/fmd/download/yc_historical.zip RBA: http://www.rba.gov.au/statistics/tables/xls/f17hist.xls BoE: http://www.bankofengland.co.uk/statistics/Documents/yieldcurve/uknom05_mdaily.xls BoC: http://www.bankofcanada.ca/rates/interest-rates/bond-yield-curves/.	−0.0002	0.0001	0.0470	−0.0145	9.7703	−0.4113	0.3947	−0.2300
$\Delta(\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^{\mathrm{EU}})$		−0.0004	0.0000	0.0784	0.4713	10.6002	−0.3974	0.9412	−0.1841
$\Delta(\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^{\mathrm{EU}})$		−0.0002	0.0000	0.0859	−0.2833	10.1404	−0.9713	0.6268	−0.1508
$\Delta(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^{\mathrm{JP}})$		−0.0007	−0.0003	0.0483	−0.1844	9.3525	−0.4275	0.3579	−0.1102
$\Delta(\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^{\mathrm{JP}})$		−0.0009	0.0005	0.0876	−0.0429	8.6251	−0.7107	0.6840	−0.1101
$\Delta(\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^{\mathrm{JP}})$		−0.0006	−0.0012	0.0891	0.0801	13.8899	−0.9807	0.8900	−0.1089
$\Delta(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^{\mathrm{GB}})$		−0.0001	0.0000	0.0440	−0.4554	10.1648	−0.4537	0.2788	−0.1825
$\Delta(\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^{\mathrm{GB}})$		0.0003	0.0010	0.0659	0.0280	7.9994	−0.4737	0.4746	−0.0708
$\Delta(\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^{\mathrm{GB}})$		0.0000	0.0000	0.0716	−0.0021	6.6027	−0.4346	0.4945	−0.0302
$\Delta(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^{\mathrm{AU}})$		−0.0005	0.0000	0.0605	0.1667	7.5503	−0.4581	0.4690	−0.3141
$\Delta(\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^{\mathrm{AU}})$		−0.0003	0.0001	0.0849	0.1280	6.5926	−0.4861	0.7147	−0.2466
$\Delta(\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^{\mathrm{AU}})$		−0.0002	0.0000	0.0871	0.0158	5.5323	−0.4674	0.4797	−0.0855
$\Delta(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^{\mathrm{CA}})$		−0.0001	0.0000	0.0337	−0.4586	13.2736	−0.4077	0.2628	−0.1100
$\Delta(\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^{\mathrm{CA}})$		−0.0002	0.0000	0.0699	−0.1771	12.4698	−0.8162	0.5590	−0.1052
$\Delta(\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^{\mathrm{CA}})$		−0.0001	0.0000	0.0961	−0.4087	86.6554	−1.9131	1.8418	−0.1531
$\Delta(\mathcal{L}_t^{\mathrm{US}} - \mathcal{L}_t^{\mathrm{CH}})$		−0.0003	0.0001	0.0436	−0.3609	9.9029	−0.4370	0.3168	−0.2063
$\Delta(\mathcal{S}_t^{\mathrm{US}} - \mathcal{S}_t^{\mathrm{CH}})$		−0.0003	0.0012	0.0694	−0.1668	7.4639	−0.5949	0.4574	−0.0893
$\Delta(\mathcal{C}_t^{\mathrm{US}} - \mathcal{C}_t^{\mathrm{CH}})$		−0.0003	−0.0005	0.0718	0.1339	8.0946	−0.4149	0.5983	−0.0858
r_t^{SP500}	Equity index returns: $r_t^j = 100(P_t^j/P_{t-1}^j - 1),\ \forall j = \{$DAX30, Nikke225, FTSE100, AllOrds, SPTSX, SPI$\}$. All equity price data are taken from Bloomberg. Prices P_t^j are PX_Last entries for the j^{th} equity price.	0.0196	0.0224	1.2649	0.0155	11.1647	−9.0350	11.5800	−0.0818
r_t^{DAX30}		0.0273	0.0486	1.5332	0.1312	7.7715	−8.4923	11.4020	−0.0167
$r_t^{\mathrm{Nikkei225}}$		0.0143	0.0000	1.5025	−0.2268	9.5657	−11.4064	14.1503	−0.0225
r_t^{FTSE100}		0.0109	0.0000	1.2230	−0.0011	9.2853	−8.8493	9.8388	−0.0423
r_t^{AllOrds}		0.0205	0.0247	0.9574	−0.4805	9.2634	−8.1980	5.5064	−0.0152
r_t^{SPTSX}		0.0273	0.0409	1.1395	−0.4777	11.9827	−9.3242	9.8233	−0.0130
r_t^{SPI}		0.0211	0.0259	1.1206	0.0059	9.5866	−7.2471	10.5788	0.0339
$\Delta\mathrm{VIX}_t$	Difference/change in the VIX and TED spread. VIX data are taken from Bloomberg. TED spread (TEDRATE) is obtained from the St. Louis Fed FRED2 database http://research.stlouisfed.org/fred2/. Returns on Gold, the BDI and Oil prices computed as $r_t^l = 100(P_t^l/P_{t-1}^l - 1),\ \forall l = \{$Gold, BDI, Oil$\}$.	−0.0036	−0.0400	1.6668	0.5539	21.8168	−17.3600	16.5400	−0.1226
$\Delta\mathrm{TED}_t$		−0.0001	0.0000	0.0563	0.7999	68.3461	−0.8000	0.9900	0.1340
r_t^{Gold}		0.0445	0.0448	1.1507	0.0226	10.0596	−9.0737	10.7883	−0.0099
r_t^{BDI}		0.0187	0.0000	1.8285	0.0915	9.6616	−11.3716	14.6341	0.7992
r_t^{Oil}		0.0810	0.0307	2.3666	0.1686	9.0281	−15.2542	23.7094	−0.0172

Notes: This table reports standard summary statistics of the 6 exchange rate returns used in the forecast evaluation. These are shown in the first 6 rows at the top of the table. This table shows also summary statistics of all the fundamental predictor variables that are used in the forecasting model. These are shown below the main horizontal line in the middle of the table, starting from row 7, and are grouped according to the 3 broad groups that are defined in (14), (15) and (16). Dashed horizontal lines are used to separate the three broad groups of fundamental predictors. The abbreviation ACF(1) denotes the first order autocorrelation of the series. These summary statistics cover the period from from January 4, 1999 to June 30, 2014 for all exchange rate and predictor variables, except for the yield curve factors of Japan and Switzerland, which end already in October 22, 2013.

Table 2: Summary statistics of the technical predictor variables use used by chartist forecasting agents

Currency	Technical Indicator	Mean	Median	Std.Dev	Skew	Kurt	Min	Max	ACF(1)	ACF(2)	ACF(3)	PACF(1)	PACF(2)	PACF(3)
EUR	$IMA_t^{(200)}$	0.4108	0.0000	0.4920	0.3627	1.1315	0.0000	1.0000	0.9426	0.9214	0.9011	0.9427	0.2950	0.1066
	$IMOM_t^{(130)}$	0.4150	0.0000	0.4928	0.3450	1.1191	0.0000	1.0000	0.9401	0.9220	0.9029	0.9404	0.3305	0.1172
	$RSI_t^{(14)}$	51.1616	50.7440	17.1477	0.0690	2.4367	1.1799	96.0865	0.9265	0.8569	0.7861	0.9269	−0.0119	−0.0477
JPY	$IMA_t^{(200)}$	0.4563	0.0000	0.4982	0.1754	1.0308	0.0000	1.0000	0.9190	0.8849	0.8602	0.9192	0.2602	0.1335
	$IMOM_t^{(130)}$	0.4756	0.0000	0.4995	0.0976	1.0095	0.0000	1.0000	0.9214	0.8825	0.8515	0.9216	0.2227	0.1007
	$RSI_t^{(14)}$	49.8455	50.5988	16.9567	−0.0051	2.5552	0.1479	99.6377	0.9196	0.8442	0.7685	0.9199	−0.0091	−0.0437
GBP	$IMA_t^{(200)}$	0.4578	0.0000	0.4983	0.1694	1.0287	0.0000	1.0000	0.9404	0.9207	0.8961	0.9406	0.3158	0.0658
	$IMOM_t^{(130)}$	0.4444	0.0000	0.4970	0.2236	1.0500	0.0000	1.0000	0.9317	0.9034	0.8811	0.9318	0.2692	0.1262
	$RSI_t^{(14)}$	51.1281	51.3032	16.9704	−0.0296	2.5395	1.4706	97.6821	0.9258	0.8533	0.7792	0.9266	−0.0287	−0.0495
AUD	$IMA_t^{(200)}$	0.3902	0.0000	0.4879	0.4500	1.2025	0.0000	1.0000	0.9302	0.8998	0.8778	0.9303	0.2577	0.1349
	$IMOM_t^{(130)}$	0.4229	0.0000	0.4941	0.3121	1.0974	0.0000	1.0000	0.9197	0.8936	0.8640	0.9199	0.3079	0.0851
	$RSI_t^{(14)}$	52.3302	52.8121	16.5720	−0.0715	2.4626	0.7477	95.9903	0.9210	0.8449	0.7682	0.9210	−0.0217	−0.0453
CAD	$IMA_t^{(200)}$	0.4180	0.0000	0.4933	0.3326	1.1107	0.0000	1.0000	0.9138	0.8784	0.8532	0.9139	0.2642	0.1398
	$IMOM_t^{(130)}$	0.4378	0.0000	0.4962	0.2509	1.0630	0.0000	1.0000	0.9142	0.8803	0.8609	0.9145	0.2705	0.1765
	$RSI_t^{(14)}$	51.7130	51.9048	16.1352	−0.0752	2.6570	0.4451	100.0000	0.9167	0.8383	0.7569	0.9179	−0.0116	−0.0603
CHF	$IMA_t^{(200)}$	0.3910	0.0000	0.4880	0.4468	1.1996	0.0000	1.0000	0.9271	0.8968	0.8686	0.9272	0.2662	0.0907
	$IMOM_t^{(130)}$	0.4123	0.0000	0.4923	0.3564	1.1270	0.0000	1.0000	0.9202	0.8985	0.8749	0.9203	0.3390	0.1223
	$RSI_t^{(14)}$	51.0410	50.7385	17.1793	0.0726	2.5061	1.2376	99.2857	0.9251	0.8524	0.7786	0.9255	−0.0248	−0.0474

Notes: This table reports standard summary statistics of all the technical predictor variables used by chartist forecasting agents for exchange rate data from January 4, 1999 to June 30, 2014. The three technical indicators $IMA_t^{(200)}$, $IMOM_t^{(130)}$ and $RSI_t^{(14)}$ denote, respectively the moving average indicator as defined in (19), the momentum indicator as defined in (20), and the RSI indicator as defined in (21). The abbreviation ACF(p) stands again for the p^{th} autocorrelation function. We also include the partial autocorrelation function denoted by PACF(q) for order q^{th} in this table to provide extra information about the dynamic structure of the technical indicators variables.

Table 3: One-step-ahead out-of-sample forecast evaluation results

Model	MSFE	Relative−MSFE	$R^2_{os}(\%)$	CW−statistic	p−value
			EUR		
Random Walk (RW)	0.3833	–	–	–	–
Chartist (C)	0.3842	1.0025	−0.2460	−0.2235	0.5884
Fundamentalist (F)	0.3798	0.9909	0.9128	3.6405	0.0001
Model combined (MC)	0.3792	0.9893	1.0728	3.6082	0.0002
			JPY		
Random Walk (RW)	0.4238	–	–	–	–
Chartist (C)	0.4247	1.0020	−0.2042	−0.3730	0.6454
Fundamentalist (F)	0.4198	0.9906	0.9351	3.7341	0.0001
Model combined (MC)	0.4207	0.9926	0.7438	3.2687	0.0005
			GBP		
Random Walk (RW)	0.3228	–	–	–	–
Chartist (C)	0.3234	1.0020	−0.2001	−0.4073	0.6581
Fundamentalist (F)	0.3201	0.9917	0.8267	3.3575	0.0004
Model combined (MC)	0.3196	0.9901	0.9888	3.4756	0.0003
			AUD		
Random Walk (RW)	0.7256	–	–	–	–
Chartist (C)	0.7259	1.0005	−0.0491	0.4491	0.3267
Fundamentalist (F)	0.7261	1.0007	−0.0675	1.3509	0.0884
Model combined (MC)	0.7234	0.9970	0.2988	1.6113	0.0536
			CAD		
Random Walk (RW)	0.3642	–	–	–	–
Chartist (C)	0.3655	1.0037	−0.3708	−1.6879	0.9543
Fundamentalist (F)	0.3633	0.9976	0.2353	1.6488	0.0496
Model combined (MC)	0.3633	0.9976	0.2448	1.4737	0.0703
			CHF		
Random Walk (RW)	0.4966	–	–	–	–
Chartist (C)	0.4981	1.0031	−0.3094	−0.6204	0.7325
Fundamentalist (F)	0.4904	0.9875	1.2453	4.2349	0.0000
Model combined (MC)	0.4896	0.9859	1.4145	4.2361	0.0000

Notes: This table reports the one-step ahead out-of-sample forecast evaluation results for the 6 exchange rates of interest. In column 1, the various models that are considered in the evaluation are listed. Columns 2 and 3 show the mean squared forecast errors (MSFE) and Relative−MSFEs corresponding to the various models, where the Relative−MSFE is computed by taking the values from column 2 and deflating them by each exchange rate's respective $\text{MSFE}_{(\text{RW})}$. In column 4 we report the Campbell and Thompson (2008) out-of-sample R^2 (denoted by R^2_{os}) which is computed as $1 - \text{MSFE}_{(\iota)}/\text{MSFE}_{(\text{RW})}, \forall \iota = \{\text{C}, \text{F}, \text{MC}\}$. The Clark and West (2007) CW−statistic and its corresponding p−value are given in columns 6 and 7. The out-of-sample evaluation period is from from September 12, 2001 to June 30, 2014 for all currencies, except for the JPY and CHF series, where the sample ends already in October 22, 2013.

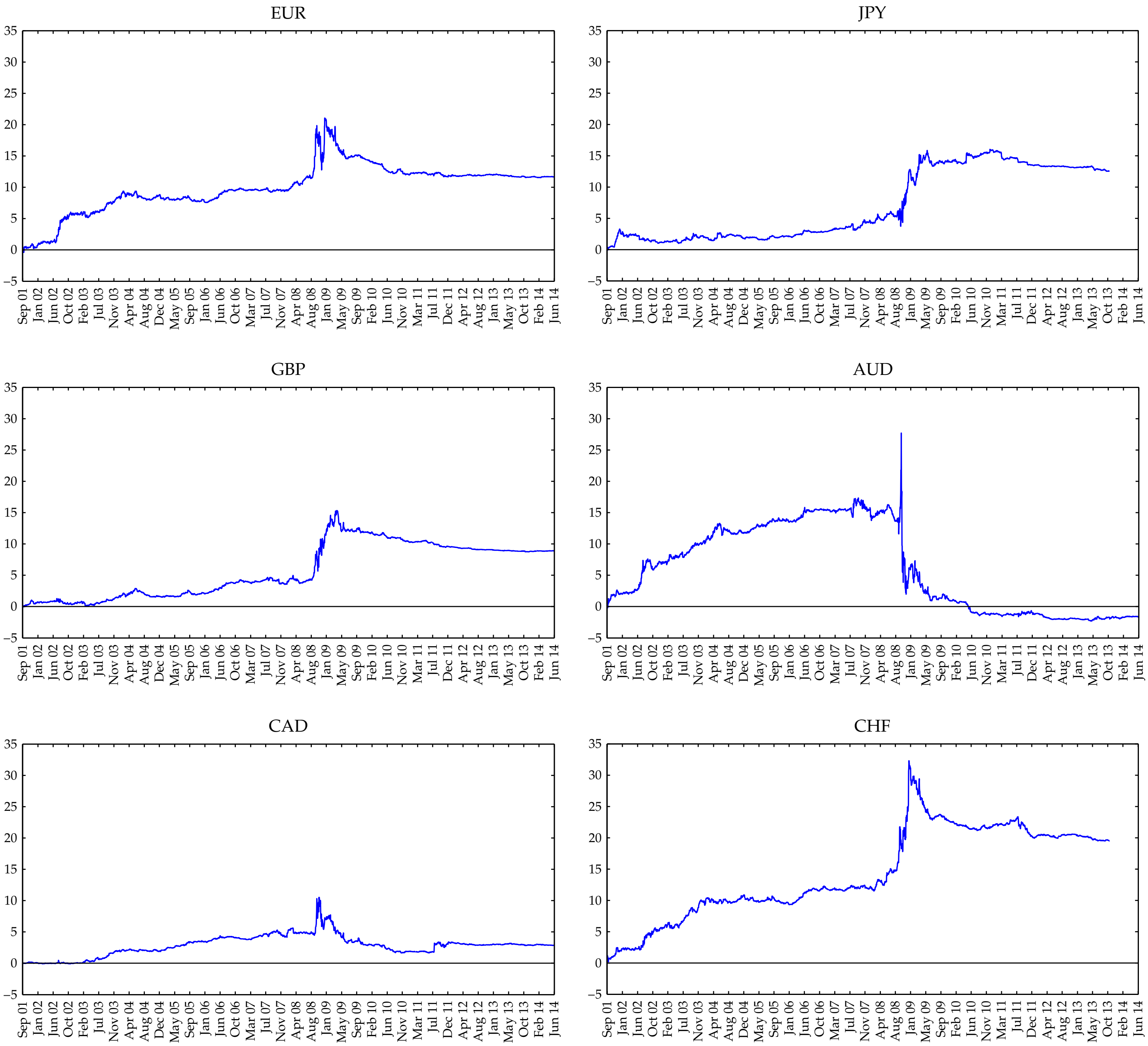

Figure 2: Time series evolution of the cumulative MSFE (cumMSFE_t) of the model combined forecasts (relative to the RW model) over the out-of-sample period from September 12, 2001 to June 30, 2014 (October 22, 2013 for the CHF and JPY series). Note that we plot all series up to June 30, 2014 so that the dates on the $x-$axis can be compared easily.

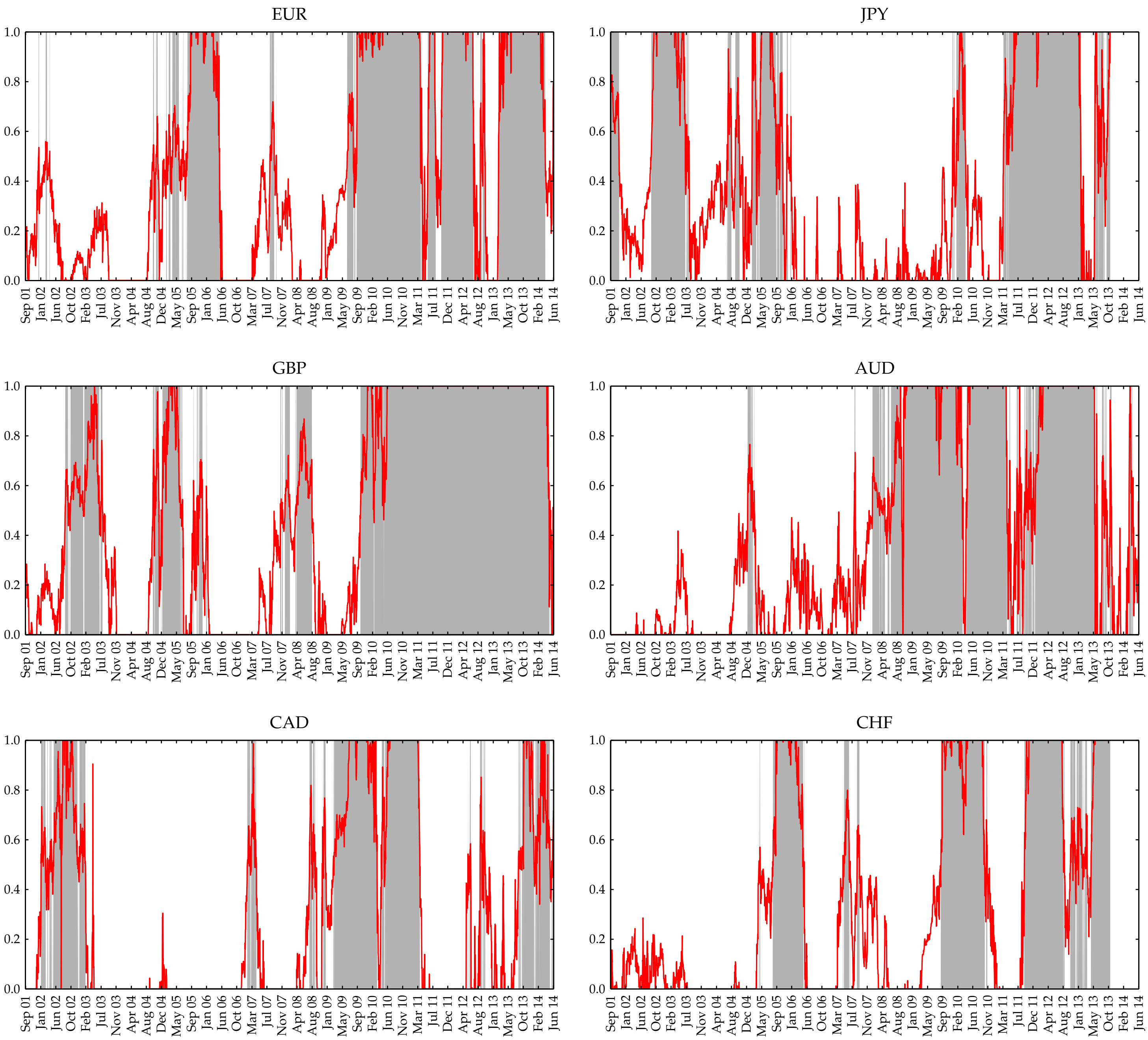

Figure 3: Time series evolution of the predicted chartist weight (or influence) function $\hat{\omega}_t$ over the out-of-sample forecasting period from September 12, 2001 to June 30, 2014. The weight function is obtained from a rolling window regression with fixed estimation window of 200 observations. That is, we get an estimate of ω_t over the sample period from December 6, 2000 to September 11, 2001 and use this first estimate as the predicted weight for September 12, 2001 and then roll through the out-of-sample observations.

Table 4: Multiple-steps-ahead out-of-sample forecast evaluation results

Forecast Horizon	$MSFE_{(RW)}$	$MSFE_{(MC)}$	Relative−MSFE	$R^2_{os}(\%)$	CW−statistic	p−value
			EUR			
$h = 2$	0.7836	0.7814	0.9973	0.2703	2.3165	0.0103
$h = 3$	1.1611	1.1576	0.9970	0.3028	1.7789	0.0376
$h = 4$	1.5236	1.5272	1.0023	−0.2335	1.2042	0.1143
$h = 5$	1.8960	1.9067	1.0056	−0.5632	0.7622	0.2230
			JPY			
$h = 2$	0.8232	0.8218	0.9982	0.1805	2.2559	0.0120
$h = 3$	1.2064	1.2089	1.0021	−0.2129	1.0190	0.1541
$h = 4$	1.5956	1.6066	1.0069	−0.6934	0.7396	0.2298
$h = 5$	1.9619	1.9739	1.0061	−0.6086	0.8519	0.1971
			GBP			
$h = 2$	0.6621	0.6612	0.9988	0.1238	1.6960	0.0449
$h = 3$	0.9929	0.9943	1.0014	−0.1384	1.0034	0.1578
$h = 4$	1.3229	1.3244	1.0012	−0.1182	0.8695	0.1923
$h = 5$	1.6489	1.6534	1.0028	−0.2756	0.5607	0.2875
			AUD			
$h = 2$	1.3894	1.3943	1.0035	−0.3532	0.7828	0.2169
$h = 3$	2.0341	2.0323	0.9991	0.0884	1.4512	0.0734
$h = 4$	2.6400	2.6400	1.0000	0.0003	1.2679	0.1024
$h = 5$	3.3067	3.3192	1.0038	−0.3774	0.9846	0.1624
			CAD			
$h = 2$	0.6686	0.6675	0.9984	0.1629	1.3548	0.0877
$h = 3$	0.9959	0.9926	0.9966	0.3378	1.8639	0.0312
$h = 4$	1.3269	1.3282	1.0010	−0.1000	0.8735	0.1912
$h = 5$	1.6835	1.6852	1.0010	−0.1038	0.8285	0.2037
			CHF			
$h = 2$	0.9596	0.9578	0.9982	0.1824	2.2605	0.0119
$h = 3$	1.4194	1.4194	1.0000	0.0008	1.3438	0.0895
$h = 4$	1.8704	1.8727	1.0012	−0.1240	1.1390	0.1273
$h = 5$	2.3065	2.3158	1.0040	−0.4028	0.6805	0.2481

Notes: This table reports the multiple-step ahead out-of-sample forecast evaluation results for the 6 exchange rates of interest. In column 1, the various forecast horizons that are considered in the evaluation are listed. Columns 2, 3 and 4 show the mean squared forecast errors (MSFEs) of the random walk (RW) benchmark, the MSFEs of the model combined (MC) predictions, and the Relative−MSFEs defined as $MSFE_{(MC)}/MSFE_{(RW)}$. In column 5 we report the Campbell and Thompson (2008) out-of-sample R^2 (denoted by R^2_{os}) which is computed as $1 - MSFE_{(MC)}/MSFE_{(RW)}$. The Clark and West (2007) CW−statistic and its corresponding p−value are given in columns 6 and 7. The out-of-sample evaluation period is from from September 27, 2001 to June 30, 2014 for all currencies, except for the JPY and CHF series, where the sample ends again in October 22, 2013. We use a HAC robust variance in the computation of the CW−statistic, employing a Quadratic Spectral Kernel on the pre-whitened cw_{t+1} series, using an ARMA$(1, 1)$ model as the approximating model and choose a data driven AR(1) bandwidth to compute the HAC of the pre-whitened series.

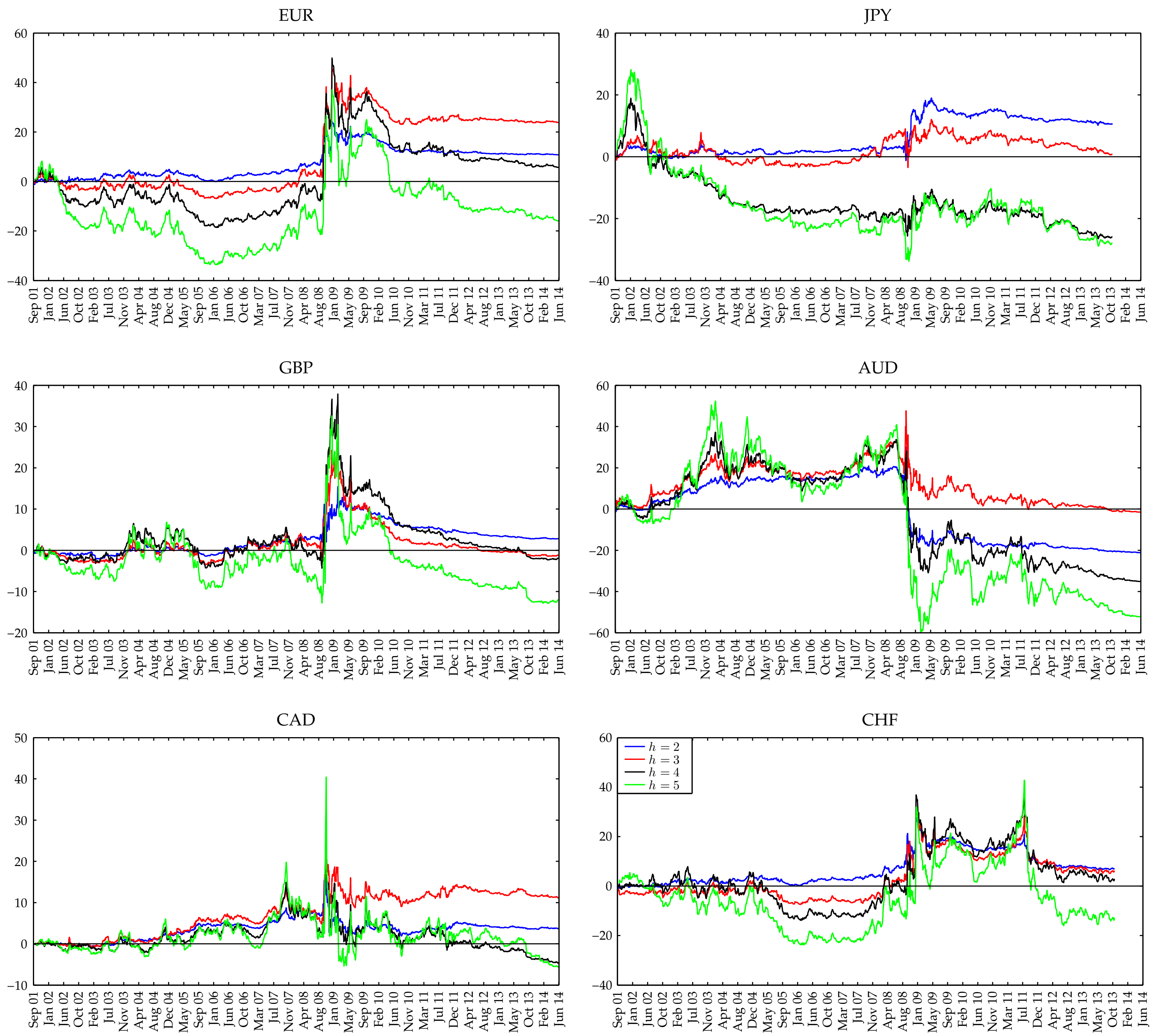

Figure 4: Time series evolution of the $h-$step ahead cumulative MSFE (cumMSFE_t) of the model combined forecasts (relative to the RW model) for forecast horizons $h = 2, 3, 4$ and 5 over the out-of-sample period from September 27, 2001 to June 30, 2014 (October 22, 2013 for the CHF and JPY series). Note that we plot all series up to June 30, 2014 so that the dates on the $x-$axis can be compared easily. Also, we off-set the forecasts for $h = 3, 4$ and 5 with 1 to three dimensional NAN vector entries in the forecasts so that the dimensions of the forecast vectors are the same and can be compared on the same time scale.