Heterogeneous Agents, the Financial Crisis and Exchange Rate Predictability

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

Gion Donat Piras
KPMG
Zürich
Switzerland

First Version: October 2, 2014
This Version (fx.6): July 8, 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 data 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 for 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. A substantial part of the forecast gains are realised around the time of the Lehman Brothers collapse from September 2008 to February 2009. Lastly, we show that our model combined forecasts produce economic value to a mean variance investor, yielding annualized Sharpe ratios of around 1.1 and performance fees in excess of 400 basis points.

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, Valentytn Panchenko, Andreas Schrimpf, George Tauchen, Graham Elliott, David Rapach, Lorenzo Camponovo, Alexander Apokin, Roland Meeks and Philip Vermeulen, as well as seminar and conference participants at the University of St. Gallen, University of Graz, the International Symposium on Econometric Theory and Applications in Tokyo (2015), the International Symposium on Forecasting in Riverside (2015), and the annual conference of the International Association for Applied Econometrics in Thessaloniki (2015) for helpful comments and discussions.

†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.
# Table of Contents

1 Introduction .................................................. 3

2 Modelling approach .......................................... 8
   2.1 Empirical model ........................................... 9
   2.2 Relation to theoretical heterogeneous agent literature ............. 10
   2.3 Constructing Fundamentalist and Chartist forecasts ................. 11

3 Data .......................................................... 16
   3.1 Fundamental variables ..................................... 18
      3.1.1 Information in the yield curve ......................... 18
      3.1.2 Information in stock prices ........................... 20
      3.1.3 Risk aversion measures and global trade activity ........ 21
   3.2 Technical variables ....................................... 23
      3.2.1 Moving Average Rules ................................. 24
      3.2.2 Momentum Indicators ................................. 25
      3.2.3 Relative Strength Index ............................... 26
   3.3 Summary statistics and visual overview ......................... 27

4 Forecast construction and evaluation ....................... 31
   4.1 Prediction setting ......................................... 31
      4.1.1 Computing individual fundamentalist and chartist forecast .......... 32
      4.1.2 Fitting and evaluation periods ........................ 34
   4.2 Evaluation criteria ....................................... 35
   4.3 Forecast evaluation results ................................ 37
      4.3.1 Statistical forecast evaluation ......................... 38
      4.3.2 Time series evolution of the model combination weights ....... 41
      4.3.3 What is driving the predictability results ............... 42
      4.3.4 Forecast performance over sub-periods .................. 45
      4.3.5 Economic evaluation of predictability ................... 47
   4.4 Forecast performance at longer horizons ................... 51
      4.4.1 Computing agent specific multiple period ahead out-of-sample forecasts 51
      4.4.2 Multiple-step ahead forecast evaluation ................ 52

5 Conclusion .................................................. 55

References ....................................................... 58

Figures and Tables ............................................. 64

Appendix ......................................................... A-1
   A.1. Calibration of $\lambda$ and $\alpha$ ............................................. A-1
1. Introduction

Since the seminal work of Meese and Rogoff (1983), it is well known that standard macro-economic 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 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 state of the macroeconomy as it is perceive by financial market participants. The financial proxy

\(^1\)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.
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. 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 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),

\[\text{See, for instance, Meese and Rogoff (1983), Mark (1995), Cheung et al. (2005), Wright (2008) for a few classic studies, and Rossi (2013) for a recent major review of the literature.}\]
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. 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. In addition to the uncertainty as to what fundamental 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. Nevertheless, despite the need

\[4\] 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 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.
for a non-linear model, empirical out-of-sample forecast evaluations in favour of non-linear exchange rate models are rather weak. 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 evaluation period from November 2001 to June 2014, we show that forecasts from the heterogeneous agent model significantly outperform the forecasts from a 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, stan-

---

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.
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. Using a dynamic asset allocation strategy we show further that the forecasts from our model combined heterogeneous agent model produce relevant ‘economic value’, yielding annualized out-of-sample Sharpe ratios of up to 1.1 and performance fees in excess of 400 basis points relative to a random walk benchmark.

The time series plots of the cumulative difference of the squared forecast errors and the cumulative portfolio wealth (relative to a random walk model) show 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 measures 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 influ-
ence 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 economically 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 had 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

This section provides a detailed description of the modelling approach that we follow. We initially describe the econometric approach that we use to build the empirical heterogeneous agent model and then provide a discussion of its links and similarities to the reduced form theoretical model of De Grauwe and Grimaldi (2005). Lastly we describe the flexible econometric framework used to construct the agent specific forecasts.
2.1. Empirical model

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 \(\mathbb{E}_t(y_{t+1}|F)\) and \(\mathbb{E}_t(y_{t+1}|C)\) denote the forecasts constructed with our fundamentalists and chartists agents’ information sets, respectively. The model combined (MC) forecast is then obtained as a weighted average of the fundamentalist and chartist predictions, that is:

\[
\mathbb{E}_t(y_{t+1}|MC) = (1 - \hat{\omega}_t)\mathbb{E}_t(y_{t+1}|F) + \hat{\omega}_t\mathbb{E}_t(y_{t+1}|C),
\]

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

\[
y_t = (1 - \omega_t)\mathbb{E}_{t-1}(y_t|F) + \omega_t\mathbb{E}_{t-1}(y_t|C) + \nu_t,
\]

\(^8\)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 $E_{t-1}(y_t|F)$ and $E_{t-1}(y_t|C)$ being respectively the forecasts of $y_t$ using information up to time $t-1$, for the two agents, and $\nu_t$ is an error term. Note here, that since we are using the time $t$ estimate of $\omega_t$ to forecast $y_{t+1}$, this effectively implies that we assume a random walk evolution for $\omega_t$. Also, in the empirical implementation of the forecast combination, we re-estimate $\omega_t$ for each new observation that becomes available using a rolling window scheme, so that a time-varying weight $\hat{\omega}_t$ 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 in the model.

### 2.2. Relation to theoretical heterogeneous agent literature

The model combined forecasts from our empirical heterogeneous agent specification defined in (1) can be viewed as the reduced form model of the theoretical heterogeneous agent models used widely in the behavioural finance literature (see Brock and Hommes (1997, 1998)). For instance, in the model of De Grauwe and Grimaldi (2005), the market expectation of the change in the exchange rate is written as a weighted average of the expectations of chartist and fundamentalists, that is, using the same notation as in De Grauwe and Grimaldi (2005), as (see equation 8 on page 696 and including transaction costs):

$$
E_t(\Delta s_{t+1}) = -\eta_{ft} \theta(s_t - s_t^*) \mathbb{I}(|s_t - s_t^*| < C) + \eta_{ct} \beta \sum_{i=0}^{T} a_i \Delta s_{t-i}
$$

(3)

where $s_t$ is the log of the exchange rate, $s_t^*$ is its fundamental value, $\mathbb{I} (\cdot)$ is an indicator variable that is equal to 1 when the statement $\mathbb{I} (|s_t - s_t^*| < C)$ is true, $E_t$ denotes the expectation operator at time $t$, $\eta_{ft}$ and $\eta_{ct}$ are, respectively, ‘profit functions’ of fundamentalist and chartist agents that are normalised to add up to unity (see equations 4 and 5 on page 694 in De Grauwe and Grimaldi (2005)). These ‘profit functions’ gauge the goodness of the individual fundamentalist
and chartist forecasting rules over time. From the relation in (3) it is evident that fundamentalists base their forecasts on deviations of the exchange rate from its fundamental value (i.e., an ‘error correction mechanism’), while chartists use a simple ‘positive feedback rule’ (or momentum), that is, they extrapolate past movements of exchange rate changes into the future.

Comparing our empirical model combined forecasting rule in (1) to the one proposed by the theoretical model of De Grauwe and Grimaldi (2005) in (3) highlights the similarities between the two. Both are constructed as weighted averages from the individual agents’ forecasts, with the combination weights being determined by a ‘goodness of performance’ measure. Nevertheless, since the forecasting environment in real time is likely to be quite complex in the sense that financial market participants tend to use a variety of fundamentalist and chartist prediction rules in practice, which are also likely to change over time, we abstract from using a single fundamentalist and/or chartist predictor variable when constructing the individual agent specific predictions and use a flexible model averaging approach instead. Exact details of the modelling approach that we use, and the intuition behind why we are using it is discussed in the next section.

2.3. 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, \( E_t(y_{t+1}|F) \) and \( E_t(y_{t+1}|C) \) 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.\(^\text{10}\) 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

\(^{10}\)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, i.e., on overbought or oversold signals.
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.\footnote{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).} 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 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 $t$.\footnote{For reasons of simplicity, we use standard $y_t$ and $x_t$ 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.} Also, let $x_{t-1}$ 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).\footnote{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 set of predictors contained in the $m^{th}$ model is denoted by $x^{(m)}_{t-1}$, with the dimension of $x^{(m)}_{t-1}$ being $(1 \times K_m)$. The two equations that make up the DMA framework (for model $m$) are:

\begin{align*}
\text{Measurement:} \quad y_t &= x^{(m)}_{t-1} \beta^{(m)}_t + u^{(m)}_t \quad \text{(4a)} \\
\text{State:} \quad \beta^{(m)}_t &= \beta^{(m)}_{t-1} + \epsilon^{(m)}_t \quad \text{(4b)}
\end{align*}

where (4a) and (4b) are measurement and state equations, respectively. The two disturbance terms $u^{(m)}_t$ and $\epsilon^{(m)}_t$ in (4) are jointly Multivariate Normal (MN) distributed, uncorrelated with each other and over time, that is:

\begin{align*}
\begin{bmatrix} u^{(m)}_t \\ \epsilon^{(m)}_t \end{bmatrix} &\sim MN \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} H^{(m)}_t & 0 \\ 0 & Q^{(m)}_t \end{bmatrix} \right) \quad \text{(5)}
\end{align*}

where $H^{(m)}_t$ and $Q^{(m)}_t$ are the variance and covariance matrix of the measurement and state
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 \( Q_t^{(m)} \) and by fixing the model set \( \mathcal{M}_t = m \), i.e., to one particular model, the system in (4) takes the form of a standard state-space model, making it thereby possible to extract or 'filter' the time varying parameters \( \beta_t^{(m)} \) 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 \( Q_t^{(m)} \), the Kalman Filter recursions are:

\[
\begin{align*}
\text{Prediction:} & \quad \hat{\beta}_{t|t-1}^{(m)} = \hat{\beta}_{t-1|t-1}^{(m)} \\
& \quad P_{t|t-1}^{(m)} = P_{t-1|t-1}^{(m)} + Q_t^{(m)} \\
& \quad \hat{y}_{t|t-1}^{(m)} = x_{t-1}^{(m)} \hat{\beta}_{t|t-1}^{(m)} \\
\text{MSE of prediction errors:} & \quad \hat{F}_t^{(m)} = x_t^{(m)} P_{t|t-1}^{(m)} x_t^{(m)\top} + H_t^{(m)} \\
\text{Kalman Gain:} & \quad G_t^{(m)} = P_{t|t-1}^{(m)} x_{t-1}^{(m)\top} / \hat{F}_t^{(m)} \\
\text{Updating:} & \quad \hat{\beta}_{t|t}^{(m)} = \hat{\beta}_{t|t-1}^{(m)} + G_t^{(m)} (y_t - \hat{y}_t^{(m)}) \\
& \quad P_{t|t}^{(m)} = P_{t|t-1}^{(m)} - G_t^{(m)} x_{t-1}^{(m)} P_{t|t-1}^{(m)} 
\end{align*}
\]

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

The Kalman Filter recursions in (6) are conditional on \( H_t^{(m)} \) and \( Q_t^{(m)} \) (and model \( m \)). To avoid having to estimate \( H_t^{(m)} \) and \( Q_t^{(m)} \), two simplifying assumptions are used in the literature. The first one, which is due to Raftery et al. (2010), is to replace \( P_{t|t-1}^{(m)} \) in (6a) by

\[
P_{t|t-1}^{(m)} = \frac{1}{\lambda} P_{t-1|t-1}^{(m)} \tag{7}
\]
where $\lambda \in [0, 1]$. This approximation implies that $Q_t^{(m)} = (\lambda^{-1} - 1) P_{t-1|t-1}^{(m)}$. In the given context, the $\lambda$ 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:

$$H_t^{(m)} = \kappa H_{t-1}^{(m)} + (1 - \kappa) \hat{u}_{t-1}^{2(m)}, \quad (8)$$

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:

$$\pi_{t|t-1}^{(m)} = \Pr(M_t = m | \mathcal{I}_{t-1}). \quad (9)$$

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

$$\hat{y}_{t|t-1}^{(DMA)} = \sum_{m=1}^{M} \hat{y}_{t|t-1}^{(m)} \pi_{t|t-1}^{(m)}, \quad (10)$$

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 (10) feasible, model prediction and updating recursions are needed. Let $p_{jm} = \Pr(M_t = m | M_{t-1} = j)$ denote the (time invariant) transi-

---

\(^{14}\)This is also known as ‘windowing’. Intuitively, we can think of $\lambda$ as a weighting function, where observations $\tau$ periods in the past receive a weight of $\lambda^\tau$. 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, Hansen and Lunde, 2005). Approximating the time varying volatility by EWMA is thus unlikely to create any important loss in accuracy. We discuss later on how the $\kappa$ parameter is calibrated.
tion probability of moving from model $j$ at time $t-1$ to model $m$ at time $t$. Also, let $f_N^{(m)}(y_{t-1} | \mathcal{I}_{t-1})$ 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 $\hat{y}_{t-1}^{(m)}$ and $F_t^{(m)}$ as computed in (6b) and (6c), respectively. That is, $f_N^{(m)}(y_{t-1} | \mathcal{I}_{t-1}) = N(\hat{y}_{t-1}^{(m)}, F_t^{(m)})$. Given an initial or prior model probability $\pi_{0|0}^{(m)}$, the model probability prediction and updating equations are then constructed as:

\[
\text{Model Probability Prediction:} \quad \pi_t^{(m)} | t-1 = \sum_{j=1}^{M} \pi_{t-1}^{(j)} P_{jm} \tag{11a}
\]

\[
\text{Model Probability Updating:} \quad \pi_t^{(m)} | t = \frac{\pi_{t-1}^{(m)} f_N^{(m)}(y_t | \mathcal{I}_{t-1})}{\sum_{j=1}^{M} \pi_{t-1}^{(j)} f_N^{(j)}(y_t | \mathcal{I}_{t-1})} \tag{11b}
\]

A final simplification that is needed to make the computation of the predictive model probabilities feasible is to approximate (11a) with

\[
\pi_t^{(m)} | t-1 = \frac{\alpha^{(m)}}{\sum_{j=1}^{M} \alpha^{(j)}},
\]

where $\alpha \in [0, 1]$. The approximation in (12) 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 $\alpha$ parameter in (12) can again be interpreted as a 'forgetting factor'.

The implementation of the DMA procedure to forecast exchange rate returns requires the calibration of the EWMA smoothing parameter $\kappa$, as well as the two forgetting factor parameters, $\lambda$ and $\alpha$. We follow the guidelines provided in RiskMetrics (1996) for daily data and fix the $\kappa$ parameter at 0.94. For the two forgetting factors, Koop and Korobilis (2012) recommend to set the values for $\lambda$ and $\alpha$ close to 1, so that the parameters (as well as the model probabilities) evolve gradually over time. We elaborate on the choice of $\lambda$ and $\alpha$ values in Section 4.

\[\text{See also Section 3.2 in Raftery et al. (2010) and pages 874 – 875 in Koop and Korobilis (2012) for additional discussion on this.}\]

\[\text{See page 97 of the documentation in RiskMetrics (1996). Note here that RiskMetrics (1996) uses $\lambda$ to denote their EWMA smoothing parameter and not $\kappa$ as we do.}\]

\[\text{See the discussion on pages 872 – 875 in Koop and Korobilis (2012). The effective window size, ie, how much of}\]
3. Data

In our empirical analysis, we use spot rates of the 6 most frequently traded currency pairs. 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, i.e., 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). 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. Our full data set consists of 4041 daily observations from the 4th of January 1999 to the 30th 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 for the Euro are available from the 4th 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 21st of October 2013.

---

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.
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. 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.

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.

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:

\[ r_t = 100\left(\frac{P_t}{P_{t-1}} - 1\right), \]  

\[ (13) \]

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.

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).

See also, Harvey (1989) or Harvey (1993) for examples of other studies that use financial data as proxies for macroeconomic variables.
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\textsuperscript{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\textsuperscript{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 (13). 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 \textit{et al.} (2003), who show that the information content in the yield curve is valuable for exchange rate forecasting.\textsuperscript{25} Moreover, in the context of macro-finance models, Diebold \textit{et al.} (2006) and Rudebusch and Wu (2008) have documented that the empirically derived level and slope factors are ‘strongly’ correlated with inflation and economic activity.\textsuperscript{26}

To capture the information in the yield curve, we construct level, slope and curvature factors, denoted by \( L_t, S_t \) and \( C_t \), using daily data on zero coupon yields. We follow Diebold \textit{et al.} (2006) and compute the three ‘empirical’ factors using linear combinations of yields of various

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

\textsuperscript{26}More specifically, Diebold \textit{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).
maturities. The level, slope and curvature factors are computed as:

\[
\text{Level : } L_t = \left( y_t^{(3)} + y_t^{(24)} + y_t^{(120)} \right) / 3
\]

(14a)

\[
\text{Slope : } S_t = (y_t^{(3)} - y_t^{(120)})
\]

(14b)

\[
\text{Curvature : } C_t = \left( 2y_t^{(24)} - y_t^{(3)} - y_t^{(120)} \right),
\]

(14c)

where \( y_t^{(\tau)} \) is the time \( t \) yield of a zero-coupon bond with maturity \( \tau \) (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. 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 for these two countries. 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

27Since 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.

28Note 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.

29We thank Andrea Vedolin for making these parameter estimates available to us. For more details on the construction of the factors and the bond data that was used, we refer to Malkhozov et al. (2014).
denoted by $x_{LSC,i}^t$ and computed as:

$$x_{LSC,i}^t = \left[ (L_{US}^t - L_i^t), (S_{US}^t - S_i^t), (C_{US}^t - C_i^t) \right],$$ (15)

where $L_{US}^t$ ($L_i^t$), $S_{US}^t$ ($S_i^t$), and $C_{US}^t$ ($C_i^t$) are level, slope and curvature factors for the US (the $i^{th}$ foreign currency), respectively, with $i = \{\text{EU, JP, GB, AU, CA, CH}\}$ being a country index for the exchange rates of interest. Due to the high persistence in the yield curve factors $x_{LSC,i}^t$, we use the (time) difference of the yield curve factors denoted by $\Delta x_{LSC,i}^t$, with $\Delta$ being the difference operator, as the predictor variables in the forecast evaluation.\(^\text{30}\)

### 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).\(^\text{31}\) 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 the UK, the SPI for Switzerland, the SPTSX for Canada and the All Ordinaries for Australia.\(^\text{32}\) 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

\(^{30}\)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 $L_{US}^t$, $S_{US}^t$ and $C_{US}^t$, 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 $(L_{US}^t - L_{EU}^t)$, $(S_{US}^t - S_{EU}^t)$ and $(C_{US}^t - C_{EU}^t)$, respectively. To avoid this high persistence, we prefer to work with the (time) differenced series for $x_{LSC}^t$, that is, $\Delta x_{LSC}^t$.

\(^{31}\)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). Also, Hatemi-J et al. (2006) show that there can be causal predictive information in stock prices, which will be relevant also for exchange rate forecasts.

\(^{32}\)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.
(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 $x_{t}^{EQT,i}$) consists of:

$$x_{t}^{EQT,i} = \left[ r_{t}^{EQT,US}, r_{t}^{EQT,i} \right],$$

where $r_{t}^{EQT,US}$ and $r_{t}^{EQT,i}$ 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 (13) 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. 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 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. 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.

---

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.

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.
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 (2012, 2014); Baumeister et al. (2015) who also use the BDI as a measure of global trade activity 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 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 $x_{t}^{RISK/ACTIV}$, includes
the following variables:

\[ x_t^{\text{RISK/ACTIV}} = [\Delta VIX_t, \Delta TED_t, r_t^{\text{GOLD}}, r_t^{\text{BDI}}, r_t^{\text{OIL}}], \]

(17)

where \( \Delta VIX_t \) and \( \Delta TED_t \) denote the (time) difference in the series of the CBOE 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 (13). 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.\textsuperscript{35}

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

\[
\begin{bmatrix}
  x_t^{\text{LSC},i} & x_t^{\text{EQT},i} & x_t^{\text{RISK/ACTIV}}
\end{bmatrix},
\]

(18)

where \( i \) denotes the foreign currency of interest.

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

\textsuperscript{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.
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).\textsuperscript{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).\textsuperscript{37}

### 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

\textsuperscript{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.

\textsuperscript{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.
prices is defined as:

\[ \text{MA}\left(^{(n)}_{t}\right) = n^{-1} \sum_{i=t-n+1}^{t} S_i, \quad \forall t \geq n \in \mathbb{Z}, \quad (19)\]

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 \( \text{MA}^{(200)}_t \) and the spot price \( S_t \).\(^{38}\)

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

\[
\text{IMA}^{(200)}_t = \begin{cases} 
1 & \text{if } S_t > \text{MA}^{(200)}_t \\
0 & \text{otherwise.}
\end{cases} \quad (20)
\]

### 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:

\[
\text{IMOM}^{(130)}_t = \begin{cases} 
1 & \text{if } S_t > S_{t-130} \\
0 & \text{otherwise.}
\end{cases} \quad (21)
\]

\(^{38}\)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.
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.\footnote{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.}

### 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 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:

$$\text{RSI}_{t}^{(n)} = 100 - \frac{100}{1 + \frac{\text{MA}_{t}^{(n)}(\text{uc}_{t})}{\text{MA}_{t}^{(n)}(\text{dc}_{t})}}$$  \hspace{1cm} (22)

where $\text{MA}_{t}^{(n)}(\xi_{t})$ denotes the $n$–period MA filter in (19) applied to variables $\xi_{t}$, and $\text{uc}_{t}$ ($\text{dc}_{t}$) are upclose (downclose) measures defined as:

$$\text{uc}_{t} = \begin{cases} 
\Delta S_{t} & \text{if } \Delta S_{t} > 0 \\
0 & \text{otherwise}
\end{cases} \quad \text{and} \quad dc_{t} = \begin{cases} 
-\Delta S_{t} & \text{if } \Delta S_{t} < 0 \\
0 & \text{otherwise,}
\end{cases}$$  \hspace{1cm} (23)
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.\footnote{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.}

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:\footnote{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.}

\begin{equation}
\left[ r_i^t, \text{IMA}_i^{(200)}, \text{IMOM}_i^{(130)}, \text{RSI}_i^{(14)} \right].
\end{equation}

### 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.

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. 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 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. 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

---

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.
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.
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. 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 summary 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. 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

---

44 For Japan and Switzerland, the level slope and curvature factors only go up to October 21, 2013.
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.
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 (20) and (21), 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.

\(^{46}\)The autocorrelation and partial-autocorrelation structure of the BDI returns (not shown here) in fact display the properties of an ARMA(1, 2) model.
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 \( \hat{\rho}^{MC}_{t+1|t} \)) are computed as:

\[
\hat{\rho}^{MC}_{t+1|t} = (1 - \hat{\omega}_t)\hat{\rho}^F_{t+1|t} + \hat{\omega}_t\hat{\rho}^C_{t+1|t},
\]

where \( \hat{\rho}^F_{t+1|t} \) and \( \hat{\rho}^C_{t+1|t} \) denote, respectively, the individual 1-step ahead forecasts constructed by fundamentalist and chartist agents at time \( t \). The (fitted) weights \( \hat{\omega}_t \) in (25) are obtained by means of a Granger and Ramanathan (1984) forecast combination regression of the form:

\[
r_t = (1 - \omega_t)\hat{\rho}^F_{t|t-1} + \omega_t\hat{\rho}^C_{t|t-1} + \nu_t
\]

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

\[
(r_t - \hat{\rho}^F_{t|t-1}) = \omega_t(\hat{\rho}^C_{t|t-1} - \hat{\rho}^F_{t|t-1}) + \nu_t,
\]

where the weights \( \omega_t \) in (27) are restricted to be in the \([0, 1]\) interval.\(^{47}\)

Two features are evident from the specification of the model averaged predictions defined in (25) and how the combination weights are obtained in (27). First, the time \( t \) estimate of \( \omega_t \)

\(^{47}\)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 \texttt{lsqlin} command in Matlab to enforce the \([0, 1]\) interval on \( \omega_{t,h} \).
is used in the averaged forecast construction in (25). This implies that the best forecast of the averaging weight \( \omega_t \) one period into the future is its time \( t \) estimate, that is, \( \mathbb{E}_t (\omega_{t+1}) = \hat{\omega}_t \). 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 2.3.

### 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 \( a = \{F, C\} \) denote the agent type for which we construct the forecast, i.e., a fundamentalist or chartist agent type. Following the general description of the DMA framework in Section 2.3, the forecasting model for (the \( i^{th} \)) exchange rate return of agent \( a \) takes the form

\[
\begin{align*}
    r_{t+1}^a &= x_t^a (m) \beta_{t+1}^a (m) + u_{t+1}^a (m) \quad (28a) \\
    \beta_{t+1}^a (m) &= \beta_t^a (m) + \epsilon_{t+1}^a (m), \quad (28b)
\end{align*}
\]

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 \( x_t^a, \forall a = \{F, C\} \) is, following (18) and (24), defined as:

\[
    x_t^F = \left[ 1, x_t^{LSC}, x_t^{EQT}, x_t^{RISK/ACTIV} \right] \quad (29)
\]

for fundamentalists and

\[
    x_t^C = \left[ 1, r_t, IMA_t^{(200)}, IMOM_t^{(130)}, RSI_t^{(14)} \right] \quad (30)
\]

for chartist agents. In (29) and (30), the intercept term is denoted by 1, and the lagged exchange rate return \( r_t \) in (30) allows for the possibility of AR type dynamics in the returns.\(^{48}\)

\(^{48}\)It should be 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 exist times when an increase in the autocorrelation of returns (in absolute terms) occurs. To allow for this possibility, we include a one period lagged exchange rate return in the chartist predictor set. This formulation is also in line with the dynamics specified in De Grauwe and Grimaldi.
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.\(^{49}\)

To compute the return forecasts for the exchange rates from the recursions in (28), time \(t\) filtered estimates \(\hat{\beta}_{t|t}^{a(m)}\) are needed to construct the optimal forecast of \(\hat{\beta}_{t+1}^{a(m)}\). Given our random walk specification of the state dynamics in (4b), this forecast is given by \(\hat{\beta}_{t|t}^{a(m)}\), for all \(m = 1, \ldots, M\). The sequence of \(\hat{\beta}_{t|t}^{a(m)}\) is obtained from the Kalman filter recursions outlined in (6). To implement the Kalman filter, we need to specify initial values. We follow Koop and Korobilis (2012) and use a diffuse prior for \(\hat{\beta}_{0|0}^{a(m)} \sim MN(0_{K_m}, 100I_{K_m})\), where \(0_{K_m}\) is a \((K_m \times 1)\) dimensional vector of zeros and \(I_{K_m}\) is a \((K_m \times K_m)\) dimensional identity matrix. The model updating probabilities for agent type \(a\) in (11b) are initialised with an uninformative prior \(\pi_{0|0}^{a(m)} = \frac{1}{M}\), so that all models are assumed to be equally likely. The \(\kappa\) term in the EWMA specification is fixed at 0.94.

The \(\alpha\) and \(\lambda\) parameters are ‘calibrated’ at 0.99 and 0.999, respectively. Exact details on how this calibration was performed are provided in the Appendix. To briefly summarise our calibration approach here, we define a two dimensional search grid over the \([0.95\, 1] \times [0.95\, 1]\) square for \(\alpha\) and \(\lambda\) and use the first 500 observations (from January 6, 1999 to December 5, 2000) as the ‘in-sample’ period. We then find the optimal \(\alpha\) and \(\lambda\) parameterisations as the values that minimize the mean squared forecast error (MSE). After we performed the grid search, it became apparent that the optimal \(\alpha\) value was always selected to be 0.99 for all currencies, and for both fundamentalist and chartist models. For \(\lambda\), the optimal value was either 0.999 or close to 0.999. That is, for 7 out of 12, the optimal \(\lambda\) value was 0.999, while for the remaining 5 it was two times 0.9995 and 0.998 respectively, and one time 0.996. To keep the calibration of the \(\lambda\) parameter as simple as possible, we preferred to fix it at the single value of \(\lambda = 0.999\) for all forecast scenarios that we consider.\(^{50}\)

\(^{49}\)Note that the null or base model is the one with only an intercept term in it.

\(^{50}\)Note here also that Koop and Korobilis (2012) and Raftery et al. (2010) have pointed out that it is common to choose a \(\lambda\) value near 1, to have a gradual evolution of the time varying coefficients in the models. Since we are using daily data, we would like to avoid having too much variation in the coefficients by using a \(\lambda\) value that is too low. 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 \(\lambda = 0.999\), observations one year ago (that is, 252 days ago) receive a weight of 0.7710 (or 77.10\%). With a lower \(\lambda\) value of say 0.99, observations one
Given the model updating probabilities \( \pi_a^{(m)}_{t+1|t} \) and \( \alpha = 0.99 \), forecasts of the model probabilities are computed as \( \pi_a^{(m)}_{t+1|t} = \frac{\pi_{a\alpha}^{(m)}_{t|t}}{\sum_{j=1}^{M} \pi_{a\alpha}^{(j)}_{t|t}} \), yielding the DMA based forecast of agent type \( a \) of the exchange rate return of interest as:

\[
\hat{r}^a_{t+1|t} = \sum_{m=1}^{M} x^a_{t} \pi_{a\alpha}^{(m)}_{t|t} \pi_{a\alpha}^{(m)}_{t+1|t}
\]

for \( a = \{F, C\} \), 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 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 require two in-sample or fitting periods. That is, we need one calibration/initialisation period to obtain the agent specific forecasts, \( \hat{r}^F_{t+1|t} \) and \( \hat{r}^C_{t+1|t'} \), and another calibration period for the combination weights \( \hat{\omega}_t \) in (25) to be determined. We use the first 500 observations (from January 6, 1999 to December 5, 2000) as the in-sample period to calibrate the \( \lambda \) and \( \alpha \) parameters, which in our Kalman Filter setting also serves as a burn-in period to minimise the influence of the initial values (or priors) on the filtered state vector \( \hat{\beta}_{t|t} \).

We then take the next 252 observations from December 6, 2000 to November 22, 2001 — which corresponds to a one year time horizon — to get our estimate \( \hat{\omega}_t \) needed to compute our first model averaged out-of-sample forecast for November 23, 2001. We then roll through the rest of the out-of-sample data to update \( \hat{\omega}_t \) using a fixed window size of 252 observations and produce (recursively updated) one-step ahead forecasts of the returns. The effective out-of-sample period thus spans from November 23, 2001 to June 30, 2014 (to October 22, 2013 for JPY and CHF), yielding 3287 (respectively 3102) evaluation points.

\footnote{year ago receive a weight of only 0.0733 (or 7.33\%). This is much too low and would imply substantial variability in the coefficients over time. We find the \( \lambda = 0.999 \) calibration to be much more reasonable. Also, our calibration of \( \alpha = 0.99 \) coincides with the recommended value in Koop and Korobilis (2012) and Raftery et al. (2010), which allows for a marginally more frequent updating of the model probabilities.}
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 $CW -$ 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 $\ell$ that we consider, where $\ell = \{RW, MC, F, C\}$. These forecast errors are computed as:

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

(32)

with corresponding MSFEs being

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

(33)

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 $\text{MSFE}_{(MC)}$ be the MSFE from the agent based model combined forecasts and let $\text{MSFE}_{(RW)}$ 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:

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

(34)

Intuitively, the $R^2_{os}$ statistic in (34) 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):

\[ CW - \text{statistic} = -\frac{2}{T_{os}} \sum_{t=T_{is}}^{T} \hat{e}_{t+1|t}^{(RW)} \left( \hat{e}_{t+1|t}^{(RW)} - \hat{e}_{t+1|t}^{(MC)} \right) \]  \hspace{1cm} (35)

(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

\[ cw_{t+1} = dm_{t+1} + adj_{t+1} \]  \hspace{1cm} (36)

where

\[ dm_{t+1} = \hat{e}_{t+1|t}^{2(RW)} - \hat{e}_{t+1|t}^{2(MC)} \]  \hspace{1cm} (37)

and

\[ adj_{t+1} = \left[ \hat{r}_{t+1|t}^{(RW)} - \hat{r}_{t+1|t}^{(MC)} \right]^2. \]  \hspace{1cm} (38)

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

\[ CW - \text{statistic} = \frac{cw}{\sqrt{\text{Var}(cw)}} \]  \hspace{1cm} (39)

where \( cw = T_{os}^{-1} \sum_{t=T_{is}}^{T} cw_{t+1} \) and \( \text{Var}(cw) \) 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.\(^{51}\)

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

\(^{51}\)See also the discussion in Section 2.1 in Diebold (2015) for more background on this in the context of the traditional Diebold-Mariano (DM) statistic.
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 $\hat{\beta}_{t|t}^{a(m)}$ for all $a = \{F, C\}$ in (31) 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 squared forecast errors (SFE) of the RW and MC forecasts over the out-of-sample period. This cumulative difference (denoted by cumSFE$_t$) is commonly used in the equity premium forecasting literature as a tool to highlight 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:

$$\text{cumSFE}_t = \sum_{t=T_{os}}^{T_{os}} \left( \hat{\varepsilon}_{t+1|t}^{2(RW)} - \hat{\varepsilon}_{t+1|t}^{2(MC)} \right).$$ (40)

A value of cumSFE$_t$ 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 cumSFE$_t$ means that the MC forecasts produce better predictions than the RW benchmark.

4.3. Forecast evaluation results

In this section, we provide a detailed statistical as well as economic analysis of the forecasts performance of our model. For readability, we break this section into 5 subsection, separately discussing the statistical evaluation results, the time varying evolution of the agent weight function, what is driving the predictability results, and a statistical evaluation over three sub-periods to measure the performance before, during and after the financial crisis. Lastly, we discuss the economic significance of our statistical results.
4.3.1. Statistical forecast evaluation

In Table 3 we present the one-step ahead out-of-sample forecast evaluation results for the period from November 23, 2001 to June 30, 2014 (to October 22, 2013 for JPY and CHF). 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) $R_{os}^2$ (in percent) as defined in (34), 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 (27).

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 1.22% for the Swiss Franc, with the Australian Dollar being the only currency with a negative $R_{os}^2$ of $-0.18\%$. From the magnitude of the CW–statistics 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 boarder line insignificant 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 $R_{os}^2$ values can be economically sizeable. 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 $R_{os}^2$ are computed at the daily frequency. To put this magnitude into perspective, in the equity premium forecasting literature, Campbell and Thompson (2008) and more recently Neely et al. (2014) have shown that $R_{os}^2$ values as low as 0.5% computed on monthly data produce eco-

\footnote{We group the results for each currency into 4 rows in Table 3.}
nomically meaningful predictive results in the sense that ‘large’ gains in portfolio performance can be obtained which an investor would be willing to pay for. Similarly, in the exchange rate forecasting literature, Della Corte and Tsiakas (2012) and Li et al. (2015) have shown that modest improvements in the out-of-sample $R^2$ relative to a benchmark random walk model can generate annual performance fees in the order of magnitude of more than 4% per year for a risk-averse investor. We return more formally to answer the question of economic significance of the model combined forecast improvements over the random walk ones in detail in Section 4.3.5.

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.\footnote{Note here also that, although chartist forecasts are inferior to fundamentalist 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 cumSFE$_t$ series, as defined in (40), 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 cumSFE$_t$ series ends already on October 22, 2013 for these two currencies. Nevertheless, for reasons of comparability

\footnote{It should be stressed here again that we do not include the observation to be forecasted in the calibration period of the weight $\omega_t$ in (25). That is, the combination weight is computed for the first 252 observations, given return forecasts from the two agent types, and the first out-of-sample forecast is then constructed for observation 253. 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.}
with respect to date entries in the figures that we show, we have plotted all series up to June 30, 2014.\(^\text{54}\) As a reminder, the cumSFE\(_t\) 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 cumSFE\(_t\) series plotted in Figure 2, one can notice the following 4 visually striking features from these plots. First, the cumSFE\(_t\) 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 cumSFE\(_t\) series is (nearly) monotonically increasing for all series up to the September 2008 period, i.e., 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.\(^\text{55}\) 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 predictions. This conclusion can be reached from the overall downward ‘trend’ that is visible in the cumSFE\(_t\) 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 cumSFE\(_t\) series is for the Australian Dollar, where the forecast accuracy of the model combined predictions initially improves dramatically, with the cumSFE\(_t\) 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 cumSFE\(_t\) had dropped below 6. For the Euro and the Canadian Dollar, similar but less accentuated swings in the cumSFE\(_t\) are visible. For the Euro, the cumSFE\(_t\) 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 cumSFE\(_t\) magnitude, also fairly abrupt movements are visible.

\(^{54}\)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.

\(^{55}\)A similar result is found in Buncic and Moretto (2015) in a forecast evaluation using LME copper data.
For instance, the cumSFE\(_t\) 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 cumSFE\(_t\) 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.\(^{56}\)

### 4.3.2. Time series evolution of the model combination weights

To learn about the influence of chartist agents, we show the time series evolution of the weight function \(\hat{\omega}_t\), as defined in (25), 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 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.

\(^{56}\)There are several papers that 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.
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. Nevertheless, it is clear that the ‘good’ performance of the fundamentalist agent based forecasts had diminished substantially since February 2009.

4.3.3. What is driving the predictability results

Given the positive forecast performance of the model combined forecasts, we can now look at the dynamic influence of the individual fundamentalist and chartist predictors over time. To do this, we examine the dynamic evolution of two quantities of interest. The first one is the posterior inclusion probability (PIP, for short) of the predictor variables. The second one is the weighted average of the updated estimates of the latent state vector, which we denote by

---

57 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.

58 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 rate differentials remained high, 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 collapse. Another key point that needs to be looked at is whether the fixed $\alpha$ and $\lambda$ 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.

---
\( \hat{\beta}_{t|t}^{(\text{DMA})} \). The \((k \times 1)\) dimensional PIP vector at time \( t \) is computed as:

\[
\text{PIP}_t = \sum_{m=1}^{M} \pi_{t|t}^{(m)} \mathbb{1}(x_t \in \mathcal{M} = m),
\]

(41)

where \( \pi_{t|t}^{(m)} \) is the updated model probability as defined in (11b) and \( \mathbb{1}(x_t \in \mathcal{M} = m) \) is an indicator variable that is equal to 1 if any of the regressors in \( x_t \) are included in the \( m^{th} \) model. Intuitively, the PIP is constructed by simply summing over all updated model probabilities \( \pi_{t|t}^{(m)} \) that contain the \( t^{th} \) predictor variable \( \{x_{i,t}\}_{i=1}^{k} \) in its model set.

The weighted average of the time varying parameter estimates is obtained as:

\[
\hat{\beta}_{t|t}^{(\text{DMA})} = \sum_{m=1}^{M} \pi_{t|t}^{(m)} \hat{\beta}_{t|t}^{(m)} \mathbb{1}(x_t \in \mathcal{M} = m),
\]

(42)

where \( \hat{\beta}_{t|t}^{(m)} \) is as defined in (6d) and \( \mathbb{1}(x_t \in \mathcal{M} = m) \) is again as above. Thus, \( \hat{\beta}_{t|t}^{(\text{DMA})} \) is computed by averaging each model’s \( \hat{\beta}_{t|t}^{(m)} \) estimate over the updated model probabilities \( \pi_{t|t}^{(m)} \) at each point in time for those regressors that are included in the model, with the averaging weight determined by \( \pi_{t|t}^{(m)} \).\(^{59}\)

These two quantities are plotted in Figures 5 to 10 for all fundamentalist and chartist predictor variables.

Each plot in Figures 5 to 10 shows the \( \hat{\beta}_{t|t}^{(\text{DMA})} \) coefficients marked by the blue solid line (left-axis scale) and the PIP\(_t\) drawn as the red solid line (right-axis scale) for all predictor variables over the entire out-of-sample period from November 23, 2001 to June 30, 2014. We intentionally set the same axes scales in all 6 plots to be able to easily compare the magnitudes of the PIP\(_t\) and \( \hat{\beta}_{t|t}^{(\text{DMA})} \) sequences across the exchange rates. There are a number of interesting features visible from these figures, which can be summarized as follows. First, the SP500 has a positive coefficient for all 6 exchange rates that we consider. What is striking here is the mild upward trending evolution in not only the model averaged parameter estimates, but also in the pos-

\(^{59}\)It is not feasible to compute the standard error of the model combined estimates, since the covariance between the parameter estimates of the different models is unknown and it is not clear how to compute it. One simplifying assumption that one could make is to set it to zero. Nevertheless, this seems unreasonable to us. One could also look at the time series evolution of the parameter estimates from the model with the full regressor set as an alternative. However, to conserve space, we do not provide these plots.
terior inclusion probabilities. Although there are some differences in the magnitudes of the PIP\(_t\) and \(\hat{\beta}_{t|t}^{(DMA)}\) series during the time of the Lehman Brothers collapse across the 6 series, all indicate that the importance of the information contained in the SP500 predictor variable increased substantially from September 2008 until about February 2009, dropping back towards its ‘trend level’ thereafter. The PIP\(_t\) for the SP500 is equal to (or very close to 1) over this period. This means that the best forecasting models, i.e., the ones with the highest predictive probability, always include the SP500 as a predictor variable.

Second, the influence of the Gold, Oil and BDI series increased also somewhat during this period for a few of the currencies. What is rather surprising to see is that the influence of the VIX and the TED spread is not as sizable during the time of the Lehman Brothers collapse as reported in other studies (i.e., Brunnermeier et al., 2009). For instance, for the AUD series, the magnitude of the TED spread coefficient jumped up (increase in magnitude) around the August/September 2007 period, when first signs of liquidity shortages were realised and interbank lending slowed down, while the magnitude of the coefficient dropped toward zero in September 2008, suggesting a diminished impact. For the Swiss Franc, on the other hand, the coefficient on the TED spread started to increase consistently from February 2010, then sharply from March 2011 until about September 2011 and did not start to decrease in magnitude until about August 2012. The VIX coefficient was largest (in absolute value) during the September 2006 to September 2007 period.

Third, the magnitude of the inclusion probabilities for quite a number of the predictor variables appears to be fairly constant at values close to 50% for most of the out-of-sample evaluation period. This indicates that these variables are included in half of the best performing models that are constructed. This also means that the weighting scheme that is used for each of the \(M = 2^k\) possible models is very close to equal weighting, i.e., \(\pi_{t|t}^{(m)} = 1/M\). One needs to be careful here to not interpret this result as an indication of un-informativeness of the variables that are included in the predictor set. Even with an equally weighted combination approach, the real benefit of model averaging is that it implies a certain type of ‘shrinkage’ estimation (see the discussion on page 845 in Rapach et al. (2010) for a specific application to forecasting the

\[60\] In a different setting, Buncic and Moretto (2015) also find that the influence of the SP500 predictor variable has increased substantially over the Lehman Brothers collapse period.
US equity premium or Elliott et al. (2013) for a broader discussion in the context of complete subset regression). As with all shrinkage type estimators, the improvement in out-of-sample forecasts performance is achieved by finding an optimal trade-off between (squared) bias and variance which define the MSFE. Thus, even predictor variables with posterior inclusion probabilities of ‘only’ around 50% contribute to an overall lowering of the MSFE and thus an improvement in the out-of-sample forecast performance.

Fourth, the information content in the chartist predictor variables is rather heterogeneous over the 6 currencies. For instance, the RSI(14) indicator is sizable for the Australian and Canadian Dollars, but less informative for the other 4 currencies, showing a mild increase in magnitude for the JPY and GBP series over the January 2012 to July 2012 period. The momentum indicator is important for the EUR and AUD series at the beginning of the out-of-sample period, while the moving average indicator increases in absolute importance during the September 2008 to January 2009 period for the Japanese Yen. The one period lagged exchange rate return has a sizeable negative coefficient from November 2003 to September 2008 for the AUD series and from February 2011 to about January 2013 for the Canadian Dollar, capturing a conditional feedback effect, ie., a reversal given an appreciation (or deprecation) of the currency.

### 4.3.4. Forecast performance over sub-periods

From the time series plots of the cumSFE, series shown in Figure 2 it is clear that there are substantial differences in the forecast performance of the model combined forecasts over time. To understand how these visual differences translate into statistical differences, we re-examine the statistical forecast evaluation results on three separate sub-periods. These are as follows: 1) from November 23, 2001 to September 1, 2008, which we call the pre Lehman Brothers collapse period; 2) from September 2, 2008 to March 1, 2009, which we consider as the period shortly before and following the Lehman Brothers collapse; and 3) from March 2, 2009 to June 30, 2014 (October 22, 2013 for the Japanese Yen and the Swiss Franc), which is our post Lehman Brothers collapse period. These results are reported in Table 5. Table 5 is organized in 3 blocks, with each

---

61 Note here also that it is well known that equally weighted forecast combinations often outperform more complicated weighting schemes in practice (see Smith and Wallis (2009) and Claeskens et al. (2015) for recent discussion of this). In a somewhat different forecasting context, Koop and Korobilis (2013, page 190) also obtain model probabilities that largely remain in the 0.3 and 0.5 interval.
block corresponding to one sub-period. Under the heading MSFE/rel. in the first column of each block we show the MSFE for the random walk model’s predictions in the ‘Random Walk (RW)’ row and the remaining 3 row entries are relative MSFEs of the chartist, fundamentalist and model combined forecasts, respectively, with the MSFE of the random walk model being in the denominator, ie., the benchmark model. The terms $R^2_{os}$, CW – stat., and $p$–value are again the Campbell and Thompson (2008) $R^2_{os}$, the Clark and West (2007) MSFE adjusted $t$–statistic and its corresponding one-sided $p$–value, as defined before.

The evaluation results over the three sub-periods in Table 5 confirm statistically the visual assessment from Figure 2. First, consider the results over the pre Lehman Brothers collapse period. The combined forecasts from our empirical heterogeneous agent model strongly reject the null hypothesis of the forecasts not being different from those of a RW model. The $p$–values corresponding to the Clark and West (2007) MSFE adjusted test are well below 0.001 for all MC forecasts. For the CHF and EUR series, the CW – statistic is over 5, while for the remaining exchange rate series it is over 3. In terms of the magnitude of the improvement, the Campbell and Thompson (2008) out-of-sample $R^2$ is as high as 1.93% and 1.83% for the EUR and CHF series. The lowest $R^2_{os}$ values, which are obtained for the JPY and GBP series, are still a sizable 0.67% and 0.79% at a daily sampling frequency. From the fundamentalist and chartist results it can be seen that the majority of the improvement in the model combined forecasts is driven largely by the fundamental predictor variables.

Looking over the results covering the Lehman Brothers collapse period shown in the second block in Table 5 one can see evidence of increased predictability (relative to the RW benchmark model) from the substantially larger out-of-sample $R^2$ values. These are as high as 8.14% for the model combined forecasts of the Swiss Franc series and sizeable 4.22%, 3.79% and 2.88% for the GBP, the EUR and the JPY. Moreover, these improvements are statistically significant at the 1% and 5% levels for the CHF and GBP, and EUR and JPY series, respectively. Note here that there are only 130 observations available to compute the CW – statistic, so that the results are (statistically) ‘seemingly weaker’ than for the longer pre Lehman Brothers collapse period, ie., the CW – statistic < 3, despite the much larger $R^2_{os}$ values. For the Canadian Dollar, the improvement is weak at only 0.77% and statistically insignificant. For the Australian Dollar,
the $R^2_{os}$ is negative, highlighting the poor out-of-sample performance of the model combined forecasts relative to the RW model during this period.

Finally, from the third sub-period covering the post Lehman Brothers collapse shown in the last results block we can see that for all exchange rates, except for the JPY series, the performance of the model combined forecasts drops off substantially relative to the RW benchmark, producing negative $R^2_{os}$ values. These results also corroborate the visual assessment from the cumSFE$_t$ series provided in Figure 2. Overall, we can conclude from the sub-period analysis that, as conjectured from the cumSFE$_t$ time series plots, there exists strong statistical evidence of exchange rate predictability before the Lehman Brothers collapse. Moreover, during the time of the Lehman Brothers collapse, predictability with respect to a benchmark RW model actually increased, resulting in much larger out-of-sample $R^2$ values.

4.3.5. Economic evaluation of predictability

How much ‘economic value’ do the improvements in the model combined forecasts produce? From the framework in Campbell and Thompson (2008) it is clear that there is a direct mapping between the increase in out-of-sample $R^2$ over the benchmark RW model and the proportional and/or absolute increase in portfolio return for a mean variance investor with one risky and one riskfree asset. For instance, from equation 14 in Campbell and Thompson (2008), we see that the proportional increase in the expected return from conditioning on a set of predictor variables is equal to

$$\left( \frac{R^2_{os}}{(1 - R^2_{os})} \right) \left( \frac{1 + SR^2}{SR^2} \right), \tag{43}$$

where $SR^2$ is the squared (per period) Sharpe ratio of the benchmark RW model.\textsuperscript{62} For small $R^2_{os}$ and $SR^2$ values, the relation in (43) is approximately equal to $R^2_{os}/SR^2$. From this it is clear that for any positive out-of-sample $R^2$ value, a proportional increase in the expected return will be realised.

In our setting, we have 6 risky assets (and one riskfree one) that need to be aggregated into a portfolio of risky assets and hence requires portfolio weights. To evaluate the economic significance of our model combined forecasts, we implement a standard dynamic asset allocation

\textsuperscript{62}For an absolute increase, see equation 13 in Campbell and Thompson (2008).
methodology, as recently used in Della Corte et al. (2009), Della Corte et al. (2011) and Li et al. (2015). The dynamic asset allocation model is implemented as follows. Consider a US based investor which aims to build a currency portfolio consisting of six currencies. Investing in US government bonds is assumed to be riskless, so that the yield on a US government bond is used as the riskfree interest rate, which we denote by $r^f_t$. The expected return on a foreign bond is equal to the expected exchange rate change plus the interest earned on the foreign bond. We take the 3 months interest rates, i.e. $y^{(3)}_t$, used in the construction of the level, slope and curvature factors as our short-term rate for both US and foreign bonds.

The dynamic asset allocation strategy that we implement and which maximizes the expected portfolio return $\hat{r}_{t+1|t}$ at each point in time is obtained by solving the following optimization problem:

$$\max_{\omega_t} \hat{r}_{t+1|t} = \omega_t^\prime \hat{r}_{t+1|t} + (1 - \omega_t^\prime) r^f_t$$

subject to $\sigma_p^* = (\omega_t^\prime \Sigma_{t+1|t} \omega_t)^{1/2}$,

where $r_{t+1}$ is a $6 \times 1$ vector of exchange rate returns, $\hat{r}_{t+1|t} = E_t(r_{t+1})$ is its respective conditional forecast formed at time $t$, $\iota$ is a $6 \times 1$ vector of ones, $\sigma_p^*$ is the target conditional volatility of the portfolio returns and $\Sigma_{t+1|t} = E_t[(r_{t+1} - \hat{r}_{t+1|t})(r_{t+1} - \hat{r}_{t+1|t})^\prime]$ is the $6 \times 6$ conditional variance-covariance matrix of $r_{t+1}$. The solution to the maximum expected return problem in (44) is given by the following $6 \times 1$ vector of risky asset weights:

$$\omega_t = \Sigma_{t+1|t}^{-1} (\hat{r}_{t+1|t} - u^f_t) \frac{\sigma_p^*}{\sqrt{K_t}}$$

$$K_t = (\hat{r}_{t+1|t} - u^f_t)^\prime \Sigma_{t+1|t}^{-1} (\hat{r}_{t+1|t} - u^f_t).$$

The realised return on the inverior’s portfolio is then computed as:

$$r^p_{t+1} = \omega_t^\prime (r_{t+1} - u^f_t) + r^f_t.$$ 

We evaluate the performance of our agent combined forecasts using the Goetzmann et al. 

---

63We closely follow the description in Della Corte et al. (2011) and Li et al. (2015) in this section.
(2007) performance measure, defined as

\[
G(r^p) = \frac{1}{1 - \gamma} \ln \left\{ \frac{1}{T_0} \sum_{t=T_0}^{T_{os}} \left( \frac{1 + r^p_{t+1}}{1 + r^f_{t+1}} \right)^{(1-\gamma)} \right\},
\]

where \( \gamma \) is the parameter that captures the investor’s relative risk aversion. The term \( G(r^p) \) is frequently interpreted as the certainty equivalent of the excess portfolio returns and does not require any assumptions regarding utility functions for a ranking of the portfolios to be constructed. We compare the performance of the agent combined forecasts to the RW benchmark by constructing the difference

\[
P = G(r^p_{MC}) - G(r^p_{RW}),
\]

where \( G(r^p_{MC}) \) and \( G(r^p_{RW}) \) denote the Goetzmann et al. (2007) performance measure computed on portfolios constructed from the agent combined forecasts and the random walk model, respectively. The performance measure \( P \) defined in (48) can be interpreted as the maximum performance fee that an investor is willing to pay to have access to the agent combined forecasts (relative to the RW benchmark). We follow the mainstream literature and report \( P \) in annualised basis points (bps). We also report standard (annualised) Sharpe ratio’s (\( SR \)s).

The economic evaluation results are reported in Table 6. We only consider the out-of-sample period for which data for all 6 exchange rates are available, that is, from November 23, 2001 to October 22, 2013. We follow Li et al. (2015) and set the relative risk aversion parameter \( \gamma \) to 6 and the target volatility of the portfolio \( \sigma^*_p \) to 10%. We further set the covariance matrix \( \Sigma_{t+1|t} \) in (44) and (45) at the unconditional covariance matrix \( (\Sigma) \) of exchange rate returns computed over the first 500 + 252 in-sample data points and do not update it. In line with the motivation given by Li et al. (2015), the reason for setting \( \Sigma_{t+1|t} = \Sigma \) is that we want to isolate the effect of the conditional mean forecasts on the portfolio performance and do not want to confound it by allowing for a time varying covariance matrix. From the results in Table 6 we can see that the model combined forecasts as well as the fundamentalist forecasts produce high economic value with (annualised) Sharpe ratios of 1.11 and 1.15, respectively. The performance fee of both the model combined and fundamentalists forecasts are over 400 annual basis points (relative to the

\[\text{For simplicity here, we assume zero trading costs.}\]
RW model), suggesting that a risk averse investor (with $\gamma = 6$) is willing to pay a fee of up to say 4% per year to have access to the model combined or fundamentalist forecasts. This finding is broadly in line with the magnitudes reported in Li et al. (2015), who use monthly data, nevertheless, relying on standard macro-economic variables that are not available in real time to produce the forecasts.

What is also interesting to see from the results in Table 6 is that the performance of chartist and random walk forecasts in terms of economic value are very similar, with the chartist’s model producing somewhat higher mean returns than the RW, albeit at the cost of generating more volatility in the portfolio returns, thereby yielding a lower Sharpe ratio of 0.8094. The model combined forecasts also produce marginally higher portfolio volatility than a simple fundamentalist model portfolio, despite the higher mean return. Note here that the random walk portfolio is constructed by using the no-change forecast of the exchange rate plus the yield on the foreign bond as the predicted return in the weight construction in (45). That means that the random walk portfolio is based on weights from a carry trade strategy (see also page 4 in Li et al. (2015) for further intuition about this result). This last fact is important to highlight as the gains in economic performance are not simply driven by a pure carry trade strategy.

In Figure 11 we plot the cumulative wealth of the dynamic out-of-sample investment strategy for a portfolio with an initial wealth of 1 US Dollar to provide a visual impression of the portfolio performance from the various forecasting models that we consider. The cumulative portfolio wealth constructed from the model combined forecasts (blue line) and also the fundamentalist forecasts only (red line) show a similar pattern to what we observed from the cumulative sum of squared forecast errors in Figure 2. For instance, there is an overall upward trend in cumulative wealth from the beginning of the out-of-sample portfolio construction period until September 2008, i.e., up to the time of the Lehman Brothers collapse. From September 2008 to about February 2009, there is a substantial increase in cumulative wealth, and from the end of February 2009, cumulative wealth remained fairly constant until the end of the sample in October 2013. The pure chartist and random walk (carry trade portfolios) perform fairly similar in the sense that at the portfolio level, the chartist weights are driven largely by the interest differential as opposed to the superior predictions from the technical trading rules. Nevertheless,
there do appear to be instances where the technical trading rules perform better than a random walk based portfolio (see for instance, the period from May 2010 to August 2011).

4.4. Forecast 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. Here we simply focus on a statistical evaluation to conserve space and avoid repetition. To construct multiple-step ahead out-of-sample forecasts from the agent based model combination, we implement the so-called ‘direct’ forecasting approach. That is, we re-formulate the relation in (25) for the general $h$-step ahead relation as:

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

(49)

where $\hat{r}_{t+h|t}^{F}$ and $\hat{r}_{t+h|t}^{C}$ 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 (27), from the regression:

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

(50)

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 (4) (again using the general $y_t$ and $x_t$ notation as in Section 2.3) as

$$
y_t = x_{t-h}^{(m)} \beta_{t,h}^{(m)} + u_t^{(m)}
$$

(51a)

$$
\beta_{t,h}^{(m)} = \beta_{t-1,h}^{(m)} + \epsilon_t^{(m)}
$$

(51b)

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.
where the \( h \) subscript in \( \beta_{t,h}^{(m)} \) signifies the relation to the \( h \)–period lagged value of \( x_t \). Using the same Kalman Filter recursions as in (6), but now on the \( h \)–period lagged relation as specified in (51) yields filtered estimates of the latent states (for each model and agent type \( a = \{ F, C \} \)), that is, \( \hat{\beta}_{t|t,h}^{a(m)} \).

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

\[
\hat{r}_{t+h|t}^a = \sum_{m=1}^{M} x_t^{a(m)} \hat{\beta}_{t+h|t}^{a(m)} \pi_{t+h|t}^{a(m)}
\]

(52a)

for all \( t = T_1, \ldots, T \), where, again due to the random walk evolution of the latent state vector \( \beta_{t+h|t}^{a(m)} \), the best forecast of \( \hat{\beta}_{t+h|t}^{a(m)} \) is its last observed filtered estimate, that is, \( E_t(\beta_{t+h|t}^{a(m)}) = \hat{\beta}_{t|t,h}^{a(m)} \).

The \( h \)–step ahead agent specific predictive model probabilities at time \( t \) are computed from

\[
\pi_{t+h|t}^{a(m)} = \frac{\pi_{t|t}^{a(x)}}{\sum_{j=1}^{M} \pi_{t|t}^{a(x)}}
\]

(53)

with \( \pi_{t|t}^{a(x)} \) being the (agent specific) filtered model probability, analogous to the definition given in (11b), \( \forall a = \{ F, C \} \).

### 4.4.2. Multiple-step ahead forecast evaluation

We use the same calibration for the \( \lambda, \alpha \) and \( \kappa \) parameters that were used in the one-step ahead prediction setting to implement the Kalman Filter recursions.\(^66\) 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 report results for the model combined forecasts for each exchange rate series at the four different forecast horizons.\(^67\) 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 \( \text{Var}(\hat{r}) \) in (39). More specifically, we follow the recommendation of Andrews and Monahan (1992) and employ a data driven band-

---

\(^66\)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.

\(^67\)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.
width 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)). 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.

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.

As was done for the one-step ahead forecasts, we again show plots of the cumSFE$_t$ 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 cumSFE$_t$ 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. From the time series plots of the cumSFE$_t$ series in Figure 4, a number of interesting features stand out. First, the increase in

68That 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 $1.3221 \sqrt{\frac{1}{T_{os}}}$, where $\hat{\alpha} = 4\hat{\rho}^2 / (1 - \hat{\rho})^4$ 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).

69That 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.
the cumSFE\(_t\) 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 cumSFE\(_t\) plot for the Australian Dollar is the improved forecast result at the 3 days ahead horizon. This is evident from the (red) cumSFE\(_t\) 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.

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 cumSFE\(_t\) 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 cumSFE\(_t\) 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 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 \( \text{cumSFE}_t \) 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, i.e., around the time of the Lehman Brothers collapse and a few months after. For the predicability results of the Swiss Franc, particularly 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 build 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 November 2001 to June 2014 we show that forecasts from our empirical heterogeneous agent model combined predictions significantly outperform the forecasts from a 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 show that these forecast gains are significant at the 10% level for the Australian and Canadian Dollars, and 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. Moreover, using a dynamic asset allocation framework we show further that our model combined forecasts also generate economic value. Annualized Sharpe ratios are as high as 1.1, yielding performance fees of over 400 basis points, relative to a random walk benchmark model.

We examine the performance of the model combined forecasts by plotting the evolution of the cumulative difference of the squared forecasts errors relative to a random walk model as well as the cumulative wealth of 1 US Dollar invested in November 2001 over time. From these plots it is evident that a consistent improvement over the random walk model is obtained from November 2001 until February 2009, with the period of the Lehman Brothers collapse showing a much strong predictive improvement as measured by the cumSFE$_t$ of the individual exchange rate forecasts as well as the cumulative wealth of a dynamically reweighted portfolio constructed from the model combined forecasts. This emphasises the changing nature of the forecast environment and warrants further investigation into the mechanisms that lead to this behaviour. From a pure forecasting perspective, what is interesting to see from the cumSFE$_t$ plots at longer forecast horizons is that the influence of the Lehman Brothers collapse remains strongly visible even at horizons of up to 5 days ahead.

Lastly, the evolution of the chartist weight function 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. A closer inspection of the time varying model averaged fundamentalist parameter estimates as well as the posterior inclusion probabilities reveals that the fundamentalist performance is largely driven by the SP500 predictor variable, whose influence consistently increased over the time of the Lehman Brothers collapse for all 6 exchange rate forecasts. Over the last third of the out-of-sample period, the influence of chartist forecasts in the MC predictions increased somewhat, 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, thereby generating portraying momentum in returns.
References


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>
<thead>
<tr>
<th>Variable Name</th>
<th>Description of all exchange rate returns and fundamental predictor variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{EU}$</td>
<td>Exchange rate returns: $r_t^{i} = 1000 (P_t^{i}/P_{t-1}^{i}) - 1$, $y_t^{i} = EUR/JPY, GBP, AUD, CAD, CHF$, where $P_t^{i}$ are FX Last entries for the $i^{th}$ currency, using the New York 5pm spot time stamp.</td>
</tr>
<tr>
<td>$r_{JPY}$</td>
<td>Differences in Level, Slope and Curvature factors, relative to the US.</td>
</tr>
<tr>
<td>$r_{AUD}$, $r_{CAD}$, $r_{CHF}$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (C_{EU} - S_{EU})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (S_{EU} - C_{EU})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (C_{JP} - S_{JP})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (S_{JP} - C_{JP})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (S_{CA} - C_{CA})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (C_{CA} - S_{CA})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (S_{GB} - C_{GB})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (C_{GB} - S_{GB})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (C_{US} - S_{US})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
<tr>
<td>$\Delta (S_{US} - C_{US})$</td>
<td>Exchange rate returns: $r_t = \frac{P_t}{P_{t-1}}$, where $P_t$ and $P_{t-1}$ are equally weighted prices of the corresponding US and foreign currencies, respectively.</td>
</tr>
</tbody>
</table>

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 the same 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 (15), (16) and (17). Dashed horizontal lines are used to separate the three broad groups of fundamental predictors. The abbreviations ACV($L$) denote the first order autocorrelation of the series. All summary statistics cover the period 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>
<thead>
<tr>
<th>Currency</th>
<th>Technical Indicator</th>
<th>Mean</th>
<th>Median</th>
<th>Std.Dev</th>
<th>Skew</th>
<th>Kurt</th>
<th>Min</th>
<th>Max</th>
<th>ACF(1)</th>
<th>ACF(2)</th>
<th>ACF(3)</th>
<th>PACF(1)</th>
<th>PACF(2)</th>
<th>PACF(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR</td>
<td>IMA(_{200}^{(200)})</td>
<td>0.4108</td>
<td>0.0000</td>
<td>0.4920</td>
<td>0.3627</td>
<td>1.1315</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9426</td>
<td>0.9214</td>
<td>0.9011</td>
<td>0.9427</td>
<td>0.2950</td>
<td>0.1066</td>
</tr>
<tr>
<td>EUR</td>
<td>IMOM(_{130}^{(130)})</td>
<td>0.4150</td>
<td>0.0000</td>
<td>0.4928</td>
<td>0.3450</td>
<td>1.1191</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9401</td>
<td>0.9220</td>
<td>0.9029</td>
<td>0.9404</td>
<td>0.3305</td>
<td>0.1172</td>
</tr>
<tr>
<td>EUR</td>
<td>RSI(_{14}^{(14)})</td>
<td>51.1616</td>
<td>50.7440</td>
<td>17.1477</td>
<td>0.0690</td>
<td>2.4367</td>
<td>1.1799</td>
<td>96.0865</td>
<td>0.9265</td>
<td>0.8569</td>
<td>0.7681</td>
<td>0.9269</td>
<td>−0.0119</td>
<td>−0.0477</td>
</tr>
<tr>
<td>JPY</td>
<td>IMA(_{200}^{(200)})</td>
<td>0.4563</td>
<td>0.0000</td>
<td>0.4982</td>
<td>0.1754</td>
<td>1.0308</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9190</td>
<td>0.8849</td>
<td>0.8602</td>
<td>0.9192</td>
<td>0.2602</td>
<td>0.1355</td>
</tr>
<tr>
<td>JPY</td>
<td>IMOM(_{130}^{(130)})</td>
<td>0.4756</td>
<td>0.0000</td>
<td>0.4995</td>
<td>0.0976</td>
<td>1.0095</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9214</td>
<td>0.8825</td>
<td>0.8515</td>
<td>0.9216</td>
<td>0.2227</td>
<td>0.1007</td>
</tr>
<tr>
<td>JPY</td>
<td>RSI(_{14}^{(14)})</td>
<td>49.8455</td>
<td>50.5988</td>
<td>16.9567</td>
<td>−0.0051</td>
<td>2.5552</td>
<td>0.1479</td>
<td>99.6377</td>
<td>0.9196</td>
<td>0.8442</td>
<td>0.7685</td>
<td>0.9404</td>
<td>−0.0091</td>
<td>−0.0477</td>
</tr>
<tr>
<td>GBP</td>
<td>IMA(_{200}^{(200)})</td>
<td>0.4578</td>
<td>0.0000</td>
<td>0.4983</td>
<td>0.1694</td>
<td>1.0287</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9404</td>
<td>0.9207</td>
<td>0.8961</td>
<td>0.9406</td>
<td>0.3158</td>
<td>0.0658</td>
</tr>
<tr>
<td>GBP</td>
<td>IMOM(_{130}^{(130)})</td>
<td>0.4444</td>
<td>0.0000</td>
<td>0.4970</td>
<td>0.2236</td>
<td>1.0500</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9317</td>
<td>0.9034</td>
<td>0.8811</td>
<td>0.9318</td>
<td>0.2692</td>
<td>0.1262</td>
</tr>
<tr>
<td>GBP</td>
<td>RSI(_{14}^{(14)})</td>
<td>51.1281</td>
<td>51.3032</td>
<td>16.9704</td>
<td>−0.0296</td>
<td>2.5395</td>
<td>1.4706</td>
<td>97.6821</td>
<td>0.9258</td>
<td>0.8533</td>
<td>0.7792</td>
<td>0.9266</td>
<td>−0.0287</td>
<td>−0.0495</td>
</tr>
<tr>
<td>AUD</td>
<td>IMA(_{200}^{(200)})</td>
<td>0.4229</td>
<td>0.0000</td>
<td>0.4941</td>
<td>0.3121</td>
<td>1.0974</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9302</td>
<td>0.8998</td>
<td>0.8778</td>
<td>0.9303</td>
<td>0.2577</td>
<td>0.1349</td>
</tr>
<tr>
<td>AUD</td>
<td>IMOM(_{130}^{(130)})</td>
<td>0.4229</td>
<td>0.0000</td>
<td>0.4941</td>
<td>0.3121</td>
<td>1.0974</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9197</td>
<td>0.8936</td>
<td>0.8640</td>
<td>0.9199</td>
<td>0.3079</td>
<td>0.0851</td>
</tr>
<tr>
<td>CAD</td>
<td>IMA(_{200}^{(200)})</td>
<td>0.4378</td>
<td>0.0000</td>
<td>0.4962</td>
<td>0.2509</td>
<td>1.0630</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9142</td>
<td>0.8803</td>
<td>0.8609</td>
<td>0.9145</td>
<td>0.2705</td>
<td>0.1765</td>
</tr>
<tr>
<td>CAD</td>
<td>IMOM(_{130}^{(130)})</td>
<td>0.4378</td>
<td>0.0000</td>
<td>0.4962</td>
<td>0.2509</td>
<td>1.0630</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9142</td>
<td>0.8803</td>
<td>0.8609</td>
<td>0.9145</td>
<td>0.2705</td>
<td>0.1765</td>
</tr>
<tr>
<td>CHF</td>
<td>IMA(_{200}^{(200)})</td>
<td>0.4180</td>
<td>0.0000</td>
<td>0.4933</td>
<td>0.3326</td>
<td>1.1107</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9138</td>
<td>0.8784</td>
<td>0.8532</td>
<td>0.9139</td>
<td>0.2642</td>
<td>0.1398</td>
</tr>
<tr>
<td>CHF</td>
<td>IMOM(_{130}^{(130)})</td>
<td>0.4378</td>
<td>0.0000</td>
<td>0.4962</td>
<td>0.2509</td>
<td>1.0630</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.9142</td>
<td>0.8803</td>
<td>0.8609</td>
<td>0.9145</td>
<td>0.2705</td>
<td>0.1765</td>
</tr>
<tr>
<td>CHF</td>
<td>RSI(_{14}^{(14)})</td>
<td>51.7130</td>
<td>51.9048</td>
<td>16.1352</td>
<td>−0.0752</td>
<td>2.6570</td>
<td>0.7477</td>
<td>95.9903</td>
<td>0.9210</td>
<td>0.8449</td>
<td>0.7682</td>
<td>0.9210</td>
<td>−0.0217</td>
<td>−0.0453</td>
</tr>
</tbody>
</table>

**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 $IMA_{200}^{(200)}$, $IMOM_{130}^{(130)}$ and $RSI_{14}^{(14)}$ denote, respectively the moving average indicator as defined in (20), the momentum indicator as defined in (21), and the RSI indicator as defined in (22). 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

<table>
<thead>
<tr>
<th>Model</th>
<th>MSFE</th>
<th>Relative−MSFE</th>
<th>$R^2_{os}$ (%)</th>
<th>CW−statistic</th>
<th>$p$−value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EUR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Walk (RW)</td>
<td>0.3839</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Chartist (C)</td>
<td>0.3847</td>
<td>1.0021</td>
<td>−0.2102</td>
<td>−0.1293</td>
<td>0.5514</td>
</tr>
<tr>
<td>Fundamentalist (F)</td>
<td>0.3805</td>
<td>0.9911</td>
<td>0.8908</td>
<td>3.5743</td>
<td>0.0002</td>
</tr>
<tr>
<td>Model combined (MC)</td>
<td>0.3796</td>
<td>0.9888</td>
<td>1.1231</td>
<td>3.6361</td>
<td>0.0001</td>
</tr>
<tr>
<td><strong>JPY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Walk (RW)</td>
<td>0.4249</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Chartist (C)</td>
<td>0.4257</td>
<td>1.0019</td>
<td>−0.1898</td>
<td>−0.3047</td>
<td>0.6197</td>
</tr>
<tr>
<td>Fundamentalist (F)</td>
<td>0.4211</td>
<td>0.9910</td>
<td>0.8990</td>
<td>3.6487</td>
<td>0.0001</td>
</tr>
<tr>
<td>Model combined (MC)</td>
<td>0.4217</td>
<td>0.9925</td>
<td>0.7504</td>
<td>3.2288</td>
<td>0.0006</td>
</tr>
<tr>
<td><strong>GBP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Walk (RW)</td>
<td>0.3239</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Chartist (C)</td>
<td>0.3244</td>
<td>1.0016</td>
<td>−0.1565</td>
<td>−0.2143</td>
<td>0.5848</td>
</tr>
<tr>
<td>Fundamentalist (F)</td>
<td>0.3213</td>
<td>0.9920</td>
<td>0.8041</td>
<td>3.2922</td>
<td>0.0005</td>
</tr>
<tr>
<td>Model combined (MC)</td>
<td>0.3209</td>
<td>0.9907</td>
<td>0.9321</td>
<td>3.3830</td>
<td>0.0004</td>
</tr>
<tr>
<td><strong>AUD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Walk (RW)</td>
<td>0.7241</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Chartist (C)</td>
<td>0.7244</td>
<td>1.0004</td>
<td>−0.0410</td>
<td>0.5078</td>
<td>0.3058</td>
</tr>
<tr>
<td>Fundamentalist (F)</td>
<td>0.7249</td>
<td>1.0012</td>
<td>−0.1176</td>
<td>1.2770</td>
<td>0.1008</td>
</tr>
<tr>
<td>Model combined (MC)</td>
<td>0.7221</td>
<td>0.9972</td>
<td>0.2777</td>
<td>1.5899</td>
<td>0.0559</td>
</tr>
<tr>
<td><strong>CAD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Walk (RW)</td>
<td>0.3662</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Chartist (C)</td>
<td>0.3676</td>
<td>1.0037</td>
<td>−0.3713</td>
<td>−1.6913</td>
<td>0.9546</td>
</tr>
<tr>
<td>Fundamentalist (F)</td>
<td>0.3654</td>
<td>0.9977</td>
<td>0.2347</td>
<td>1.6412</td>
<td>0.0504</td>
</tr>
<tr>
<td>Model combined (MC)</td>
<td>0.3655</td>
<td>0.9979</td>
<td>0.2115</td>
<td>1.3848</td>
<td>0.0831</td>
</tr>
<tr>
<td><strong>CHF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Walk (RW)</td>
<td>0.4962</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Chartist (C)</td>
<td>0.4978</td>
<td>1.0031</td>
<td>−0.3097</td>
<td>−0.6766</td>
<td>0.7507</td>
</tr>
<tr>
<td>Fundamentalist (F)</td>
<td>0.4902</td>
<td>0.9878</td>
<td>1.2157</td>
<td>4.1641</td>
<td>0.0000</td>
</tr>
<tr>
<td>Model combined (MC)</td>
<td>0.4897</td>
<td>0.9868</td>
<td>1.3209</td>
<td>4.0895</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**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 MSFE(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 $1 − \text{MSFE}_i/\text{MSFE}_{(RW)}$, $\forall i = \{C, F, MC\}$. 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 November 23, 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 sum of squared forecast errors (\text{cumSFE}_t) of the model combined forecasts (relative to the RW model) over the out-of-sample period from November 23, 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 \text{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 November 23, 2001 to June 30, 2014 (October 22, 2013 for the CHF and JPY series). The weight function is obtained from a rolling window regression with fixed estimation window of 1 year (252 daily observations). That is, we get an estimate of \( \omega_t \) over the sample period from December 6, 2000 to November 22, 2001 and use this first estimate as the predicted weight for November 23, 2001 and then roll through the out-of-sample observations.
<table>
<thead>
<tr>
<th>Forecast Horizon</th>
<th>MSFE_{(RW)}</th>
<th>MSFE_{(MC)}</th>
<th>Relative−MSFE</th>
<th>$R^2_{os}$ (%)</th>
<th>CW−statistic</th>
<th>$p$−value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.7836</td>
<td>0.7814</td>
<td>0.9973</td>
<td>0.2703</td>
<td>2.3165</td>
<td>0.0103</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>1.1611</td>
<td>1.1576</td>
<td>0.9970</td>
<td>0.3028</td>
<td>1.7789</td>
<td>0.0376</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>1.5236</td>
<td>1.5272</td>
<td>1.0023</td>
<td>−0.2335</td>
<td>1.2042</td>
<td>0.1143</td>
</tr>
<tr>
<td>$h = 5$</td>
<td>1.8960</td>
<td>1.9067</td>
<td>1.0056</td>
<td>−0.5632</td>
<td>0.7622</td>
<td>0.2230</td>
</tr>
<tr>
<td>JPY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.8232</td>
<td>0.8218</td>
<td>0.9982</td>
<td>0.1805</td>
<td>2.2559</td>
<td>0.0120</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>1.2064</td>
<td>1.2089</td>
<td>1.0021</td>
<td>−0.2129</td>
<td>1.0190</td>
<td>0.1541</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>1.5956</td>
<td>1.6066</td>
<td>1.0069</td>
<td>−0.6934</td>
<td>0.7396</td>
<td>0.2298</td>
</tr>
<tr>
<td>$h = 5$</td>
<td>1.9619</td>
<td>1.9739</td>
<td>1.0061</td>
<td>−0.6086</td>
<td>0.8519</td>
<td>0.1971</td>
</tr>
<tr>
<td>GBP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.6621</td>
<td>0.6612</td>
<td>0.9988</td>
<td>0.1238</td>
<td>1.6960</td>
<td>0.0449</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>0.9929</td>
<td>0.9943</td>
<td>1.0014</td>
<td>−0.1384</td>
<td>1.0034</td>
<td>0.1578</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>1.3229</td>
<td>1.3244</td>
<td>1.0012</td>
<td>−0.1182</td>
<td>0.8695</td>
<td>0.1923</td>
</tr>
<tr>
<td>$h = 5$</td>
<td>1.6489</td>
<td>1.6534</td>
<td>1.0028</td>
<td>−0.2756</td>
<td>0.5607</td>
<td>0.2875</td>
</tr>
<tr>
<td>AUD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 2$</td>
<td>1.3894</td>
<td>1.3943</td>
<td>1.0035</td>
<td>−0.3532</td>
<td>0.7828</td>
<td>0.2169</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>2.0341</td>
<td>2.0323</td>
<td>0.9991</td>
<td>0.0884</td>
<td>1.4512</td>
<td>0.0734</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>2.6400</td>
<td>2.6400</td>
<td>1.0000</td>
<td>0.0003</td>
<td>1.2679</td>
<td>0.1024</td>
</tr>
<tr>
<td>$h = 5$</td>
<td>3.3067</td>
<td>3.3192</td>
<td>1.0038</td>
<td>−0.3774</td>
<td>0.9846</td>
<td>0.1624</td>
</tr>
<tr>
<td>CAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.6686</td>
<td>0.6675</td>
<td>0.9984</td>
<td>0.1629</td>
<td>1.3548</td>
<td>0.0877</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>0.9959</td>
<td>0.9926</td>
<td>0.9966</td>
<td>0.3378</td>
<td>1.8639</td>
<td>0.0312</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>1.3269</td>
<td>1.3282</td>
<td>1.0010</td>
<td>−0.1000</td>
<td>0.8735</td>
<td>0.1912</td>
</tr>
<tr>
<td>$h = 5$</td>
<td>1.6835</td>
<td>1.6852</td>
<td>1.0010</td>
<td>−0.1038</td>
<td>0.8285</td>
<td>0.2037</td>
</tr>
<tr>
<td>CHF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.9596</td>
<td>0.9578</td>
<td>0.9982</td>
<td>0.1824</td>
<td>2.2605</td>
<td>0.0119</td>
</tr>
<tr>
<td>$h = 3$</td>
<td>1.4194</td>
<td>1.4194</td>
<td>1.0000</td>
<td>0.0008</td>
<td>1.3438</td>
<td>0.0895</td>
</tr>
<tr>
<td>$h = 4$</td>
<td>1.8704</td>
<td>1.8727</td>
<td>1.0012</td>
<td>−0.1240</td>
<td>1.1390</td>
<td>0.1273</td>
</tr>
<tr>
<td>$h = 5$</td>
<td>2.3065</td>
<td>2.3158</td>
<td>1.0040</td>
<td>−0.4028</td>
<td>0.6805</td>
<td>0.2481</td>
</tr>
</tbody>
</table>

Notes: This table reports 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 November 26, 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 squared forecast errors ($\text{cumSFE}_t$) 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 November 26, 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.
Figure 5: Time series evolution of the (DMA) model averaged $\beta_{it}$ estimates (blue solid line, left-axes scale) and the posterior inclusion probabilities (PIPS) (red solid line, right-axes scale) for the EUR series over the out-of-sample forecast evaluation period from November 23, 2001 to June 30, 2014. The plots show the parameter estimates of all the fundamentalist (plots 1 to 10) as well as the chartist predictor variables (plots 11 to 14), respectively.
Figure 6: Time series evolution of the (DMA) model averaged $\beta_{1t}$ estimates (blue solid line, left-axes scale) and the posterior inclusion probabilities (PIPS) (red solid line, right-axes scale) for the JPY series over the out-of-sample forecast evaluation period from November 23, 2001 to June 30, 2014. The plots show the parameter estimates of all the fundamentalist (plots 1 to 10) as well as the chartist predictor variables (plots 11 to 14), respectively. The fundamentalist based forecasts end on October 22, 2013 (due to the yield curve data).
Figure 7: Time series evolution of the (DMA) model averaged $\beta_{1,t}$ estimates (blue solid line, left-axes scale) and the posterior inclusion probabilities (PIPS) (red solid line, right-axes scale) for the GBP series over the out-of-sample forecast evaluation period from November 23, 2001 to June 30, 2014. The plots show the parameter estimates of all the fundamentalist (plots 1 to 10) as well as the chartist predictor variables (plots 11 to 14), respectively.
Figure 8: Time series evolution of the (DMA) model averaged $\beta_{it}$ estimates (blue solid line, left-axes scale) and the posterior inclusion probabilities (PIPS) (red solid line, right-axes scale) for the AUD series over the out-of-sample forecast evaluation period from November 23, 2001 to June 30, 2014. The plots show the parameter estimates of all the fundamentalist (plots 1 to 10) as well as the chartist predictor variables (plots 11 to 14), respectively.
Figure 9: Time series evolution of the (DMA) model averaged $\beta_{it}$ estimates (blue solid line, left-axes scale) and the posterior inclusion probabilities (PIPS) (red solid line, right-axes scale) for the CAD series over the out-of-sample forecast evaluation period from November 23, 2001 to June 30, 2014. The plots show the parameter estimates of all the fundamentalist (plots 1 to 10) as well as the chartist predictor variables (plots 11 to 14), respectively.
Figure 10: Time series evolution of the (DMA) model averaged $\beta_{t|t}$ estimates (blue solid line, left-axes scale) and the posterior inclusion probabilities (PIPS) (red solid line, right-axes scale) for the CHF series over the out-of-sample forecast evaluation period from November 23, 2001 to June 30, 2014. The plots show the parameter estimates of all the fundamentalist (plots 1 to 10) as well as the chartist predictor variables (plots 11 to 14), respectively. The fundamentalist based forecasts end on October 22, 2013 (due to the yield curve data).
The table reports one-step ahead out-of-sample forecast evaluation results for the 6 exchange rates of interest over three different sub-periods. These are: 1) pre-Lehman Brothers collapse from November 23, 2001 to September 1, 2008; 2) Lehman Brothers collapse from September 2, 2008 to March 1, 2009 and 3) post-Lehman Brothers collapse from March 2, 2009 to June 30, 2014 (October 22, 2013 for CHF and JPY). In the column under the MSFE/rel. heading (in each sub-sample block) we show the MSFE of the random walk model in the row corresponding to Random Walk (RW), followed by the relative MSFEs of the chartist, fundamentalist and model combined forecasts, respectively, with the MSFE of the random walk being the denominator. The terms $R^2_{os}$, CW-statistic and $p$-value are as defined in Table 3.
**Table 6:** Economic evaluation of the out-of-sample forecasts

<table>
<thead>
<tr>
<th>Investment Strategy</th>
<th>$\mu_p$ (%)</th>
<th>$\sigma_p$ (%)</th>
<th>$SR$</th>
<th>$P$ (bps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Walk (RW)</td>
<td>8.0630</td>
<td>9.3230</td>
<td>0.8649</td>
<td>–</td>
</tr>
<tr>
<td>Chartist (C)</td>
<td>9.4325</td>
<td>11.6532</td>
<td>0.8094</td>
<td>53</td>
</tr>
<tr>
<td>Fundamentalist (F)</td>
<td>19.1128</td>
<td>16.6270</td>
<td>1.1495</td>
<td>422</td>
</tr>
<tr>
<td>Model Combined (MC)</td>
<td>20.3030</td>
<td>18.2361</td>
<td>1.1133</td>
<td>467</td>
</tr>
</tbody>
</table>

**Notes:** This table reports the economic value of the various forecasting models that are considered. These are constructed over the out-of-sample period from November 23, 2001 to October 22, 2013 for all six currencies that we consider. We build maximum expected return portfolios, using a relative risk aversion parameter $\gamma$ of 6 and a target portfolio volatility level of $\sigma_p^* = 10\%$. We use the first 500+252 observations from January 6, 1999 to November 22, 2001 to compute the variance covariance matrix of the exchange rate returns $\Sigma$ needed to construct the portfolio weights in (45) and do not update $\Sigma$ as we roll through the out-of-sample period. In columns two to five above, we show the annualised mean $\mu_p$ (in percent), volatility $\sigma_p$ (in percent), Sharpe ratio $SR$ and performance fee $P$ (in basis points, bps), respectively.
Figure 11: Cumulative portfolio wealth for the different forecasting models. Initial wealth is set to 1 US Dollar. The cumulative wealth is constructed from the dynamic investment strategy using the out-of-sample forecasts to construct the portfolio weights. The out-of-sample forecasts are constructed from the model combined (blue line), fundamentalist (red line), chartist (green line) and random walk model (black dashed line), respectively, over the out-of-sample period from November 23, 2001 to October 22, 2013. The horizontal line at 1 marks the initial wealth of 1 US Dollar.
Appendix for:
‘Heterogeneous Agents, the Financial Crisis and Exchange Rate Predictability’

This appendix provides some extra results to supplement the findings in the main part of the paper.

A.1. Calibration of $\lambda$ and $\alpha$

We calibrate the $\lambda$ and $\alpha$ parameters by minimising the mean squared forecast errors (MSE) via grid search on the first 500 observations, that is, on data from January 6, 1999 to December 5, 2000. We use the following grid values for $\lambda$ and $\alpha$, respectively. 

$\alpha = [0.95 \ 0.99 \ 0.995 \ 0.999 \ 1]$ and 

$\lambda = [0.95:0.01:0.99 \ 0.995:0.001:0.999 \ 0.9995 \ 1]$.

The notation $[a:c:b]$ is standard Matlab notation for: increment $a$ by $c$ up to $b$. Note here that performing a grid search over $\lambda$ and $\alpha$ is computationally demanding. We therefore do this only once, i.e., on the first 500 ‘in-sample’ observations. The ‘estimates’ of $\lambda$ and $\alpha$ are not updated as we roll through the sample. Also, due to the extra computational burden, we do not calibrate the $\kappa$ parameter on the our in-sample data, but simply fix $\kappa$ at 0.94, which is the recommended value by RiskMetrics (1996) for daily data.

Notice from the $\lambda$ and $\alpha$ grids that we specify a somewhat coarser grid for $\alpha$, which is the forgetting factor used in the construction of the predictive model probabilities (see equation (12) in the model description section). As can be seen from the grid search plots below, there is not much variation in the MSE when it comes to different values of $\alpha$. We limit our search over the $[0.95 \ 1]$ interval, which is the recommended range set by Koop and Korobilis (2012). For the $\lambda$ parameter, there can be more variation at times and we thus find it important to use a somewhat finer search grid when considering values close to one.

In Figure A.1 and Figure A.2 we show the results from the grid search. In each figure, the vertical axis shows the MSE and the horizontal axis (the $x-$axis) the $\lambda$ grid values. In each plot there are five differently coloured solid lines corresponding to the 5 $\alpha$ grid points that we use. The dotted horizontal and vertical lines mark the minimum MSE value and the corresponding $\lambda$ value at minimum, respectively. The numeric $\lambda$ and $\alpha$ values at min MSE after the grid search are also shown in the plots. From the plots in Figure A.1 and Figure A.2 it is clear that for all 6 exchange rate pairs, irrespective of whether fundamentalist or chartist predictor variables are use, the ‘best’ $\alpha$ value is attained at $\alpha^* = 0.99$. For 7 out of the 12 plots, the ‘best’ $\lambda$ value is found at $\lambda^* = 0.999$. The remaining 5 are found at values close to 0.999, that is, two are at 0.9995 and 0.998 respectively, and one is 0.996. Given these grid search based optimal values, we calibrate for reasons of simplicity all fundamentalist and chartist DMA parameters at $\lambda = 0.999$ and $\alpha = 0.99$ and do not vary them for the 6 different exchange rates and the various forecast horizons that we consider.
Figure A.1: Grid search over $\lambda$ and $\alpha$ ($\kappa = 0.94$): Fundamentalist model calibration results.
Figure A.2: Grid search over $\lambda$ and $\alpha$ ($\kappa = 0.94$): Chartist model calibration results.