Forecasting Exchange Rates: An Iterated Combination Constrained Predictor Approach

Antonios K. Alexandridis | Kent Business School | CeQuFin

*Working Paper No. 003 – October 2018*
Forecasting Exchange Rates: An Iterated Combination
Constrained Predictor Approach

Antonios K. Alexandridis* Ekaterini Panopoulou Ioannis Souropanis
Kent Business School, University of Kent, United Kingdom

October 15, 2018

Abstract

Forecasting exchange rate returns is of great interest to both academics and practitioners. In this study we forecast daily exchange rate returns of six widely traded currencies using combination and dimensionality reduction methods. We propose a hybrid Iterated Combination with Constrained Predictor approach. In addition, we examine the impact of positivity constraints on the forecasting ability of each method. Our results indicate that the proposed hybrid method outperforms the simple linear bivariate method and both the Iterated Combination and the Predictor Constrained approaches. Positivity constraints significantly improve the forecasting ability of all methods.

JEL classification: C52, C53, F31

Keywords: Exchange rates, Forecasting, Dimension reduction methods, Forecast combinations, Constrained predictors

*Corresponding author. Antonios Alexandridis, Kent Business School, University of Kent, Canterbury CT2 7FS, United Kingdom. Tel.: 0044 1227 827702, Email: A.Alexandridis@kent.ac.uk.
1 Introduction

Exchange rate (FX) forecasting is a hot topic of discussion in academic literature. In their seminal paper, Meese and Rogoff (1983) have risen the issue of no predictability in exchange rate. This gave rise to a voluminous literature on the so-called exchange rate disconnect puzzle. Mark (1995) overturned this finding in favor of FX predictability. In a recent review, Rossi (2013) argues that exchange rate predictability is affected by several factors, such as the model under consideration and forecast horizon to name a few.

In light of this puzzle, academia has turned to more sophisticated techniques, which stimulated research and led to more promising results. There are several studies which discuss and implement both linear and nonlinear approaches in forecasting FX and other asset classes, as well. Nonlinear approaches mainly belong to the machine learning literature and include neural networks, genetic programming, support vector machines and related hybrid models. 1 Recently, we observe a backward shift in the literature in favor of simpler/standard models in order to predict exchange rates. For a more detailed discussion, see among others, Orphanides (2003, 2008), Molodtsova and Papell (2009, 2012), Rossi (2013), Byrne, Korobilis and Ribeiro (2016, 2018) and Beckmann and Schüssler (2016).

This study investigates whether daily exchange rates can be forecasted with the use of financial variables. We propose a novel forecasting approach by creating a hybrid model combining two recently developed state-of-the-art methodologies. The first is an extension of the simple combination approach and was proposed by Lin, Wu and Zhou (2017). This methodology combines the forecasts with those of the benchmark by attributing weights according to their past performance and creates an iterated combination forecast. The second is also a recently introduced methodology, proposed by Pan, Pettenuzzo and Wang (2018). The authors set constraints directly to the predictors in order to take advantage of extreme shifts in the information set that may have an actual impact on the forecasting process. Our approach, namely the Iterated Combination Constrained Predictor (ICCP) approach, generates forecasts by transforming the predictor series in line with the Constrained Predictors approach and then applying the Iterated Combination approach. Our set of specifications ranges from simple univariate models including one predictor at a time to dimensionality reduction and forecast combination techniques. Specifically, we employ simple combination methods, Principal Component Analysis

---

(PCA) (see, among others Neely, Rapach, Tu and Zhou, 2014) and Partial Least Squares (PLS) (see Kelly and Pruitt; 2013, 2015). In order to incorporate the reluctance of an investor to bet on a negative return forecast, we also consider the Campbell and Thompson (2008) framework and truncate negative return forecasts at zero.

Our dataset consists of daily observations of six widely traded currencies for the period extending from February 02, 1999 to December 31, 2017 with the out-of-sample period starting on January 1st, 2004. The exchange rates we consider are the British Sterling (GBP), Japanese Yen (YEN), Swiss Franc (CHF), Euro (EUR), Canadian Dollar (CAD) and Australian Dollar (AUD) against the US Dollar (USD). We employ predictors that contain different types of information that approximate macroeconomic and financial conditions on a daily basis. Our set of 14 candidate predictors is related to risk aversion (Buncic and Piras, 2016), global trading and activity (Calvet, Fisher and Thompson, 2006; Ferraro, Rogoff and Rossi, 2015; Baumeister, Guerin and Kilian, 2015), yield curve data (Buncic and Piras, 2016), national stock indices (Christiansen, Schmeling and Schrimpf; 2012) and their respective trading volumes.

Our findings suggest that the proposed hybrid ICCP approach can actually deliver superior forecasts. Moreover, positivity constraints in the forecasts significantly improve the forecasting ability of all predictors and combination or dimensionality reduction methods for all approaches. Our robustness checks, which include the change of the length of the control window, the frequency of the data and the out-of-sample period further verify the superiority of our forecasting approach.

The remainder of the paper is organized as follows. Section 2 presents our forecasting approaches. Section 3 describes our dataset and Section 4 presents the empirical findings. In Section 5, we present the economic evaluation approach and findings. Section 6 presents the robustness checks, while Section 7 concludes the paper by providing a brief summary of our results.

2 Methodology

Our aim is to forecast exchange rate returns for six widely traded currencies. We employ the information contained in 14 financial variables in order to generate forecasts on a daily basis. First we compute the daily log returns of the exchange rates. Under the bivariate framework, we initially test each predictor individually. Then we apply combination forecasts and dimensionality reduction techniques.

For each of the predictive variables we estimate the following bivariate model:

\[ r_{t+1} = a + bx_t + u_{t+1} \]

where \( r_{t+1} \) is the daily exchange rate log return, \( b \) is the slope coefficient, \( x_t \) is the predictor under consideration and \( u_{t+1} \) is the disturbance term. Hence, we generate the daily out-of-
sample forecasts for each predictor with the use of simple OLS, such as:

\[ \hat{r}_{t+1} = \hat{a} + \hat{b}x_t \]

We denote as \( C^0_i \) the simple bivariate model for each individual predictor \( i \).

2.1 Forecast Combination Approaches

2.1.1 Iterated Combination

We apply the Iterated Combination (IC) approach in the context of exchange rate returns forecasting. This method was recently introduced by Lin, Wu and Zhou (2017) in the context of corporate bond returns forecasting. The proposed methodology is an extension of existing combination approaches. The final forecast is a weighted average of the generated forecast and the benchmark; in our case, the Random Walk (RW) with drift:

\[ r_{t+1} = (1 - \delta)\bar{r}_t + \delta \hat{r}_{t+1} + u_{t+1} \]

where \( \delta \) is the weight, \( \bar{r}_t \) is the RW forecast and \( \hat{r}_{t+1} \) is the forecast of the individual predictor. The closer \( \delta \) is to zero, the less information is contained in the candidate forecasting model. The values of \( \delta \) are estimated by minimizing the in-sample squared error, so that:

\[ \delta = \frac{\text{cov}_t(r_{t+1} - \bar{r}_t, \hat{r}_{t+1}^{\text{fitted}} - \bar{r}_t)}{\text{var}_t(\hat{r}_{t+1} - \bar{r}_t)} \]

where \( \delta \) is iteratively computed, \( \bar{r}_t \) is the sample mean of \( r_t \) using all observations until time \( t \) (RW) and \( \hat{r}_{t+1}^{\text{fitted}} \) are the fitted values (in-sample) of the individual predictor. Then, the iterated combination forecasts are calculated by:

\[ \hat{r}_{IC}^{t+1} = (1 - \hat{\delta})\bar{r}_t + \hat{\delta} \hat{r}_{t+1} \]

where the process is iterated until the end of sample.

We denote this method as \( C^{{IC,0}}_i \) for each predictor \( i \).

2.1.2 Constrained Predictor

The second method that we adapt in our setting is the Constrained Predictor (CP) method that sets constraints directly to the predictors and was proposed by Pan, Pettenuzzo and Wang (2018). Following the notation of the original paper, the predictors are transformed according to the following relation:

\[ x^*_i(n) = \begin{cases} 
  x_t & \text{if } x_t > \max(x_{t-1}, x_{t-2}, \ldots, x_{t-n}) \text{ or } x_t < \min(x_{t-1}, x_{t-2}, \ldots, x_{t-n}) \\
  0 & \text{otherwise} 
\end{cases} \]
where \( n \) is the "look-back" period. In this study we apply a 25 day control window roughly corresponding to a one-month trading period. Hence, a constrained out-of-sample forecast is generated by:

\[
\hat{r}_{t+1}^*(n) = \hat{a}^*(n) + \hat{b}^*(n)x_t^*(n) \tag{6}
\]

where \( \hat{r}_{t+1}^*(n) \) is the constrained out-of-sample forecast for \( t + 1 \) of the individual predictor and \( \hat{a}^*(n) \) and \( \hat{b}^*(n) \) are the estimated parameters of the regression. The procedure is repeated for each out-of-sample step. A drawback of this method is that only a few periods with abnormal behavior on the predictors have an actual impact on the model. To alleviate this, Pan, Pettenuzzo and Wang (2018) include the information of "normal" periods and propose a revised constrained forecast, \( \hat{r}_{t+1}^{CP}(n) \), as follows:

\[
\hat{r}_{t+1}^{CP}(n) = 0.5\hat{r}_t + 0.5\hat{r}_{t+1}^*(n) \tag{7}
\]

We denote this method as \( C_{i}^{CP,0} \) for each predictor \( i \).

2.1.3 Constrained Predictor with Iterated Combination Forecasts

In this study we propose a new method in the context of exchange rate forecasting. The proposed method is a hybrid approach of the IC and the PC methods; namely the Iterated Combination Constrained Predictor (ICCP) approach. Forecasts are generated in three steps. Initially, we constrain the predictors following relationship (5). Next we apply the IC methodology on both the constrained and unconstrained predictors. Last, we generate the forecasts using equations (6) and (7). We denote the hybrid method as \( C_{i}^{ICCP,0} \) for each predictor \( i \).

2.1.4 Forecast Combination Approach

We also apply a forecast combination approach (see Timmermann, 2006; De Zwart, Markwat, Swinkels and van Dijk, 2009; Rapach, Strauss and Zhou, 2010; Beckmann and Schüssler, 2016; Buncic and Piras, 2016). We generate forecasts on the basis of each individual predictor and then combine the individual forecasts using a simple average. The general formula for combining \( N \) individual forecasts is given by:

\[
\hat{r}_{t+1}^{POOL} = \sum_{i=1}^{N} w_i \hat{r}_{i,t+1} \tag{8}
\]

We assume an equal weight for each predictor, i.e. \( w_i = \frac{1}{N} \) where in our case \( N = 14 \). Despite its simplicity, the naive combination of forecasts is widely used in the literature. In this framework, we take advantage of the aforementioned diversification by simply merging all forecasts and calculating a simple average. We denote this method as \( C_{i}^{POOL,0} \). When IC, CP or the hybrid ICCP approach is applied the notation changes to \( C_{i}^{IC,0} \), \( C_{i}^{CP,0} \) or \( C_{i}^{ICCP,0} \),
respectively.

2.2 Dimensionality Reduction Techniques

Our dataset consists of a large number of predictors. Hence, dimensionality reduction techniques may help us extract the relevant information from the dataset. In this study we apply the Principal Component Analysis (PCA) and the Partial Least Squares (PLS) methods to transform our large set of variables to a few new predictors by extracting all relevant information.

2.2.1 Principal Components

Following, among others, Neely, Rapach, Tu and Zhou (2014), we use Principal Component Analysis (PCA) in order to reduce the dimensionality of the data and model complexity. PCA decreases the large number of predictors by transforming closely related variables to new uncorrelated ones that capture maximum variability. The predictive variables.

The daily out-of-sample forecast at time $t + 1$ obtained from the principal components is denoted as $\hat{\gamma}_{t+1}^{PCA}$ and is given by the following formula:

$$\hat{\gamma}_{t+1}^{PCA} = \hat{a} + \sum_{k=1}^{K} \hat{b}_k \hat{F}_{k,t}$$

where $\hat{F}_{k,t}$ is the $k$-th principal component estimated at time $t$.

By construction, most of the available information is concentrated in the first few components. We take into account at most the first $K = 4$ principal components, i.e. $\hat{F}_t = (\hat{F}_{1,t}, ..., \hat{F}_{K,t}), K = 1, ..., 4$. The regression parameters $\hat{b}_k$ are recursively calculated with the OLS method. The optimal number of principal components is chosen using the adjusted $R^2$ of the in-sample period.

We denote this method as $C^0_{PCA}$. When IC, CP or the hybrid ICCP approach is applied, the notation changes to $C^0_{IPA}, C^0_{CPA}$ or $C^0_{ICCP}$, respectively.

2.2.2 Partial Least Squares

A method closely related to PCA and multiple linear regression is Partial Least Squares (PLS), introduced by Wold (1966) and more recently successfully extended and adopted in finance by Kelly and Pruitt (2013, 2015). The methodology is applicable to problems with extensive datasets and demonstrates promising results (see, for instance, Stivers 2018). Contrary to PCA, PLS takes into account the relationship between the dependent and independent variables by explaining the maximum variation in the target variable. Hence, theoretically, PLS is superior to PCA.

We follow Stivers (2018) and apply the de Jong (1993) SIMPLS algorithm to extract one target relevant factor ($z_t$) from the set of potential predictors. The daily out-of-sample forecast
at time $t+1$ obtained from the PLS is denoted as $\hat{r}_{t+1}^{PLS}$ and is given by the following formula:

$$\hat{r}_{t+1}^{PLS} = \hat{a} + \hat{b}_1 z_t$$

(10)

We denote this method as $C_{0}^{PLS}$. When IC, CP or the hybrid ICCP approach is applied the notation changes to $C_{IC}^{0PLS}$, $C_{CP}^{0PLS}$ or $C_{ICCP}^{0PLS}$, respectively.

2.3 Positivity Constraints

In this last part of the experiment, we follow a growing part of literature supporting different types of constraints (see among others Ang and Piazzesi, 2003; Campbell and Thompson, 2008; and Pettenuzzo, Timmermann and Valcanov, 2014). We adopt the approach proposed by Campbell and Thompson (2008) and truncate the forecasts of returns at zero if the forecast is negative. The intuition behind this truncation is that investors are not interested in negative returns. The forecasts are transformed under the following positivity constraint:

$$\hat{r}_t^+ = \begin{cases} \hat{r}_t & \text{if } \hat{r}_t > 0 \\ 0 & \text{otherwise} \end{cases}$$

(11)

Our notation changes to $C_{i}^{+}$, $C_{i}^{IC,+}$, $C_{i}^{CP,+}$ or $C_{i}^{ICCP,+}$ for each predictor/ model specification $i$.

3 Dataset

In this study we forecast six widely traded currencies; the British Sterling (GBP), Japanese Yen (YEN), Swiss Franc (CHF), Eurozone’s Euro (EUR), Canadian Dollar (CAD) and Australian Dollar (AUD) against the US Dollar (USD). FX spot prices were collected from Bloomberg database. Our sample contains daily observations that extend from February 02, 1999 to December 31, 2017. The total number of observations is 4,940. The first 25 observations of the sample serve as a control window (“look-back period”), in order to generate the constrained predictors, as illustrated in equation (5). The following 1,257 observations are considered as the in-sample period and the remaining 3,658 are used as the out-of-sample period.

Panel A of Figure 1 shows the daily spot exchange rates of the six currencies under consideration against the US dollar for the period under examination (February 02, 1999 to December 31, 2017). Similarly, Panel B of Figure 1 depicts the daily returns of the six currencies. In Table 1 the descriptive statistics of the returns of the six currencies are presented. The mean for all currencies is similar and almost zero. On the other hand, significant differences are observed in the standard deviation, the skewness and the kurtosis. More precisely, the lowest standard deviations are observed in the GBP and CAD returns (0.58 and 0.56 respectively), while the largest ones in CHF and AUD (0.73 and 0.80 respectively). Small negative skewness is observed in the cases of YEN, (-0.09) and EUR, (-0.05), while positive ones for GPB, (0.83), CAD, (0.11),
and AUD, (0.35). Finally, larger negative skewness is observed in the case of CHF. Similarly, CHF exhibits very large kurtosis, 112.13. The kurtosis in the remaining currencies range from 4.45 in the case of EUR to 14.20 in the case of GBP.

Our set of predictors contains three groups of financial variables which can be viewed as proxies for the state of the economy. These candidate predictors are related to risk aversion and global trading/ economic activity, stock market data and yield curve data.

More in detail, we employ the VIX (CBOE) and the TED spread in order to gain a measure of ‘risk aversion’ in the markets. VIX measures the volatility implied by option prices on the S&P500 and can gauge investors’ expectations about stock market volatility over the next month. The TED spread is calculated as the difference between the 3-month LIBOR rate and the 3-month Treasury Bill rate and is related to credit/liquidity risk in the US economy. In general, an increase in VIX and/or the TED spread is associated with negative financial outlook. We also consider gold returns (GOLD), which is a safe haven against shocks in risky assets and complements VIX and TED as risk measures (Capie, Mills and Wood; 2005). Following Calvet, Fisher and Thompson (2006) and Ferraro, Rogoff and Rossi (2015), we employ the returns of Crude Oil (OIL) which is closely linked with the macroeconomic environment via inflation and changes in the interest rates. To proxy for trade activity and future demand (Baumeister, Guerin and Kilian, 2015), we employ the Baltic Dry Index returns (BDI), which is composed of four indices, the Baltic Capesize, Panamax, Handysize and Supramax, illustrating shipping activity. Finally, we examine the predictive performance of the Commodity Research Bureau (CRB) Index, which is particularly important for commodity export/import countries. CRB is the arithmetic average of the futures prices of 19 commodities and is structured as follows; 39% of the commodities are related to energy, 41% to agriculture, 7% to precious metals and the remaining to base/industrial metals.

The next set of predictors is related to equity markets, which contain information about the macroeconomic outlook of the countries considered. Specifically, we include the returns of the MSCI global index, which represents large and mid-cap equity performance across 23 developed markets countries. We also take into account the information embedded in the returns and trading volume of S&P500 (denoted as SP500 and VSP500) and the leading equity market indices of the respective currencies (denoted as EquityM and VEquityM). The equity indices we employ are FTSE100, NIKKEI225, SPI, DAX30, SPTSX and AllOrds for GBP, YEN, CHF, EUR, CAD and AUD, respectively.

Following Bekaert and Hodrick (1992), Clarida, Sarno, Taylor and Valente (2003) and Buncic and Piras (2016) we include a set of predictors related to the yield curve. We construct the level, slope and curvature factors (see Diebold, Rudebusch and Aruoba, 2006), denoted as $L_t$, $S_t$ and $C_t$ as a combination of yields of zero coupon bonds with various maturities, denoted as...
\( y_t^{(m)} \) as follows:

\[
egin{align*}
L_t &= \left( y_t^{(3)} + y_t^{(24)} + y_t^{(120)} \right) / 3 \\
S_t &= \left( y_t^{(3)} - y_t^{(120)} \right) \\
C_t &= \left( 2y_t^{(24)} - y_t^{(3)} - y_t^{(120)} \right)
\end{align*}
\]

(12)

Candidate predictors are generated by taking the differences between the respective factor for the US and each country under consideration, so that:

\[
\begin{align*}
\Delta L &= \Delta(L^{US} - L^i) \\
\Delta S &= \Delta(S^{US} - S^i) \\
\Delta C &= \Delta(C^{US} - C^i)
\end{align*}
\]

for \( i = \{GBP, YEN, CHF, EUR, CAD, AUD\} \). Table 2 provides a more detailed description of the predictors we employ, their construction and the respective data sources.

4 Out-of-sample Performance

In order to evaluate the out-of-sample forecasting performance of candidate FX return models we split our dataset into two parts. The first part is called in-sample and is used for the fitting of our models. The second part is called out-of-sample and is used for the evaluation of the proposed models. The in-sample part ranges over 1,282 values (5 years), 25 of which correspond to the control window. The out-of-sample dataset consists of the remaining 3,658 values. We produce 1-day-ahead out-of-sample forecasts recursively, i.e. the in-sample dataset is expanding at each time \( t \). We only use data up to time \( t \) in order to forecast the FX returns at the next day, \( t + 1 \).

The forecasting accuracy is measured by the Mean Square Forecasting Error (MSFE). The MSFE of each proposed model is compared against the MSFE of the random walk (RW) with drift, that is the historical average. This benchmark has proven very difficult to outperform and is calculated as follows:

\[
\hat{r}_{t+1}^{RW} = \frac{1}{T} \sum_{t=1}^{T} r_t
\]

We evaluate the forecasts of the proposed specifications \( j \) over the benchmark by calculating
the Campbell and Thompson (2008) out-of-sample $R^2$, denoted as $R^2_{OOS}$, which is given by:

$$R^2_{j,oos} = 1 - \frac{MSFE_j}{MSFE_{RW}}$$

We can interpret $R^2_{OOS}$ as the proportional change in the MSFE of the competing predictor $j$ against the MSFE obtained by the benchmark. When $R^2_{OOS}$ is positive the model under consideration generates superior forecasts than the benchmark, and vice versa.

To test for the statistical significance of positive $R^2_{OOS}$ we employ the adjusted MSFE, $MSFE_{adj}$, proposed by Clark and West (2007). The test is computed as follows:

$$MSFE_{adj} = \frac{1}{P} \left\{ \sum_{t=R+1}^{T-1} \left\{ (r_{t+1} - \hat{r}_{t+1}^{RW})^2 - [(r_{t+1} - \hat{r}_{t+1})^2 - (\hat{r}_{t+1}^{RW} - \hat{r}_{t+1})^2] \right\} \right\}$$

where $P$ is the number of out-of-sample observations, $T$ is the number of the total sample size, $r_{t+1}$ is the actual return at time $t + 1$, $\hat{r}_{t+1}^{RW}$ is the forecast generated by the benchmark and $\hat{r}_{t+1}$ is the forecast of candidate models. $MSFE_{adj}$ is composed of two terms, the first one is the MSFE of the parsimonious model and the second one is composed of the MSFE of the extended model and the average squared difference between the forecasts of the parsimonious model and those of the extended model. $MSFE_{adj}$ is an one-sided test where $H_0$ is given by $MSFE_{RW} \leq MSFE_j$ against the alternative. The standard normal distribution can provide a very good approximation of the critical values.

### 4.1 Empirical Findings

In this section, we evaluate the out-of-sample performance of the forecasting methods considered in section 2. Initially we examine whether individual predictors or combinations can provide consistently superior forecasts, irrespective of the currency under consideration. Secondly, we examine whether the proposed methodologies, namely IC, CP and ICCP, enhance the forecasting performance. Finally, we examine whether the application of positivity constraints on forecasts further improves forecast accuracy.

Table 3 shows the out-of-sample performance of all predictors and all proposed methods for the six currencies. By examining Table 3 we can make the following general observations: 1) Applying positivity constrains improves the forecasting ability of each method, 2) $C^{ICCP,0}$ outperforms the alternative unconstrained specifications $C^0$, $C^{IC,0}$ and $C^{CP,0}$, 3) $C^{ICCP,+}$ produces the highest $R^2_{OOS}$ in most cases, outperforming all other methods, 4) it is very difficult to forecast the returns of some currencies, and 5) for each currency we can identify the predictors that outperform the remaining ones, however, their performance is not constant across all currencies.

Focusing on GBP in Panel A of Table 3 we observe that SP500, MSCI and the curvature yield...
curve factor outperform the remaining predictors. Combining the forecast of each individual predictor (POOL forecast) improves the $R^2_{OOS}$ (0.12%). On the other hand, the $R^2_{OOS}$ of both PCA and PLS are negative. Applying the Iterated Combination approach we observe an increase in terms of $R^2_{OOS}$, however, only MSCI and SP500 have positive and statistically significant $R^2_{OOS}$. $R^2_{OOS}$ further increases for most predictors when the Predictor Constrained method is used. For example, the MSCI increased from 0.05% in the $C^0$ to 0.10% in the $C^{IC,0}$ and to 0.17% in the $C^{ICP,0}$. Nevertheless, the overall performance is relatively poor with only 6 predictors with positive $R^2_{OOS}$. We observe the odd fact that negative $R^2_{OOS}$ values (PLS) are accompanied by statistically significant CW critical values. Finally, when ICCP is used, $R^2_{OOS}$ further increases in most cases, e.g. the $R^2_{OOS}$ for the SP500 increases from 0.10% to 0.28%. However, again in most cases the $R^2_{OOS}$ is negative. A closer inspection of Table 3 reveals that the application of positivity constraints significantly improves the forecasting ability of each method. For $C^+$ there are only two negative $R^2_{OOS}$ while in the case of $C^{IC,+}$ and $C^{ICCP,+}$ all $R^2_{OOS}$ are positive. For example we observe that $\Delta VIX$ returns a very low -0.32% $R^2_{OOS}$ in the case of $C^0$ while it increases to a statistically significant 0.10% for $C^+$ and to a statistical significant 0.17% for $C^{ICCP,+}$. Similar changes are observed for all individuals predictors. Furthermore, the dimensionality reduction techniques (PCA and PLS) are now positive and statistically significant for all methods. In general $C^{ICCP,+}$ outperforms alternative methods when positivity constraints are applied followed by $C^+$, $C^{CP,+}$ and $C^{IC,+}$.

Focusing on YEN in Panel B of Table 3 we observe mixed results for the forecasting ability of each predictor and method. In general MSCI, SP500, OIL, $\Delta VIX$, and the slope of the yield curve deliver positive and statistically significant $R^2_{OOS}$ while on the other hand the $R^2_{OOS}$ from $\Delta TED$, GOLD, BDI, VSP500 and $\Delta L$ is negative and in general large in absolute values. Finally, $R^2_{OOS}$ for POOL, PCA and PLS are 0.24%, 0.40% and 0.20%, respectively and statistically significant at the 1% level. In general, we observe that $C^{ICCP,0}$ outperforms the alternative methods while comparing $C^0$, $C^{IC,0}$ and $C^{CP,0}$ we get mixed results. Applying positivity constraints improves the performance of all four methods. For example, $R^2_{OOS}$ for CRB increases from 0.04% to 0.16%. On the other hand, we observe that the $R^2_{OOS}$ of POOL, PCA and PLS decrease when the positivity constraint is applied although it is still positive and statistically significant. For example, in the case of POOL, $R^2_{OOS}$ decreases from 0.40% to 0.20%. Again, $C^{ICCP,+}$ outperforms the alternative methods followed by $C^{CP,+}, C^+$ and $C^{IC,+}$. Finally, it is worth noting that PCA outperforms all other specifications for all methods followed by MSCI, SP500, POOL, PLS and OIL.

Moving to CHF we observe poor forecasting ability from all predictors with an exception

\[TABLE 3 AROUND HERE\]

\[2This can be expected as CW tests for equal performance in the population, while $R^2_{OOS}$ presents the performance in a finite sample.\]
of $\Delta S$, which delivers a statistically significant $R^2_{OOS}$ of 0.12%. As previously, $C^{ICCP0}$ outperforms all unconstrained methods followed by $C^{IC0}$. On the other hand $C^{CP0}$ and $C^0$ show poor forecasting ability. Applying positivity constraints significantly improves the forecasting accuracy of all methods. For example, out of 14 individual predictors and 3 combination methods, only 2 have a positive $R^2_{OOS}$ in the case of $C^0$ while this number rises to 10 in the case of $C^+$. Similarly, we observe 6 positive $R^2_{OOS}$ in the case of $C^{ICCP0}$ while there are 13 in the case of $C^{ICCP+}$. Observing Panel C of Table 3 we conclude that ICCP with positivity constraints produces the best results, the IC approach outperforms the CP approach while the simple bivariate models rank last.

Focusing on the EUR we observe similar results to CHF. A closer inspection of Panel D of Table 3 reveals that only the volume of DAX, the slope and POOL have a positive $R^2_{OOS}$ for $C^0$. However, only the volume of the equity market (DAX) is statistically significant. We observe similar poor performance for $C^{CP0}$. When we focus on $C^{IC0}$ and $C^{ICCP0}$ we clearly observe more positive $R^2_{OOS}$, however statistical significance is obtained only in the case of OIL for $C^{IC0}$ and $\Delta TED$ and OIL for $C^{ICCP0}$. Positivity constrains significantly improve the results for all methods and the majority of predictors have a positive $R^2_{OOS}$. The proposed hybrid $C^{ICCP+}$ approach clearly outperforms all other methods obtaining the highest $R^2_{OOS}$ in 13 cases out of 17. Furthermore, the $R^2_{OOS}$ of 9 predictors is statistically significant, i.e. $\Delta VIX$, $\Delta TED$, GOLD, OIL, CRB, VSP500, $\Delta L$, $\Delta S$ and $\Delta C$. Comparing the simple bivariate model with the proposed advanced method, it is clear that significant gains are obtained, e.g. for $\Delta VIX$ the $R^2_{OOS}$ increased from -0.88% for $C^0$ to 0.11% for $C^{ICCP+}$. Finally, POOL is constantly positive for all methods, while PCA and PLS are always negative.

Next, we focus on the Canadian Dollar. The results are presented in Panel E of Table 3. In all unconstrained specifications we observe very few positive $R^2_{OOS}$ values. It is worth mentioning that GOLD outperforms the benchmark across all methods. $C^{ICCP0}$ outperforms the remaining unconstrained methods. It is also clear that positivity constraints significantly improve the forecasting power of all methods. In the cases of $C^{IC+}$ and $C^{ICCP+}$ all models, with an exception of PLS, outperform the benchmark. The results of $C^{CP+}$ and $C^+$ are similar where the majority of predictors have a positive $R^2_{OOS}$. The proposed hybrid $C^{ICCP+}$ approach ranks first followed by $C^{CP+}$, $C^{IC+}$ and $C^+$. It is noteworthy that by combining the iterated combinations and predictor constrained approach yields better results compared to each approach separately. For example in the case of SP500 the $R^2_{OOS}$ is $-0.25\%$, 0.14% and -0.11% for $C^{IC0}$ and $C^{CP0}$ respectively while it is 0.08% for the $C^{ICCP0}$ and 0.12% for $C^{ICCP+}$. As previously, we also find that POOL always outperforms both dimensionality reduction approaches.

Finally, we examine the performance of the predictors and the proposed approaches in forecasting Australian Dollar returns. The results are presented in Panel F of Table 3. In general, for the unconstrained specifications all predictors have poor performance across all
methods, nevertheless, ICCP outperforms all other methods. Positivity constraints improve
the results however when we focus on \( C^{CP+} \), \( C^{IC+} \) and \( C^+ \) we observe only 7, 10, 6 positive \( R^2_{OOS} \), while in the case of \( C^{ICCP+} \) there are 12 positive \( R^2_{OOS} \). In general, we can conclude that
\( C^{ICCP+} \) significantly outperforms all other methods, however, it is clear that is more difficult
to forecast the returns of the Australian Dollar FX returns than other currencies. Furthermore,
POOL, PCA and PLS have a poor performance although POOL outperforms both PCA and
PLS.

Overall, the results presented in Table 3 show evidence of predictability in daily exchange
rate returns. The performance of individual predictors is not similar across currencies and
methods. Predictors may outperform the benchmark in one currency however they may not
generate consistently good forecasts for every currency. Hence, the investor faces a predictor
selection problem. For this reason, we construct three set of forecasts that take into account
information from all predictors, i.e. POOL, PCA and PLS. Among the three, we observe that
the \( R^2_{OOS} \) of POOL is larger for all currencies under consideration. On the other hand, when
positivity constraints are applied we observe that OIL, the volume of S&P 500 (VSP500) and
\( \Delta C \) have constantly positive \( R^2_{OOS} \) for all currencies and methods. Similarly, GOLD and POOL
have always positive \( R^2_{OOS} \) when positivity constraints are applied for all methods and currencies
except for AUD. In general, positivity constraints significantly improve the forecasting ability
of all methods. Finally, the proposed hybrid approach that combines iterated combination and
constrained predictor outperforms all methods for all currencies both in the unconstrained and
positivity constrained setting.

We complement our analysis by examining the evolution of the cumulative squared error
difference between the benchmark and the proposed methods. Our results are presented in
Figure 2. Due to space limitations, we only present the relative figures for the combination and
dimensionality reduction techniques. In Panel (a) and (b) the results for GPB are presented.
More precisely, Panel (a) presents the results of the POOL, PCA and PLS techniques for GBP
using the \( C^0 \) method while the \( C^{ICCP+} \) is used in Panel (b). Panel (a) reveals that POOL is
always above the benchmark while PCA and PLS show negative cumulative results at the end
of the period. On the other hand, in Panel (b) we observe that all three methods are always
positive, however the performance of POOL is superior to PCA and PLS as the slope is almost
consistently positive. The same results are observed in Panel (c) for YEN. On the other hand,
the superiority of POOL is clear in panels (d)-(g). POOL is almost always positive and always
outperforms the benchmark at the end of the period for all currencies. It is worth mentioning
that cumulative squared difference between the benchmark and the POOL is almost always
positive even in the case of AUD which is the most difficult to predict currency. On the other
hand both PCA and PLS start positive but they perform worse than the benchmark for CHF,
EUR, CAD and AUD.
5 Economic Evaluation

In this section we examine the economic value of our forecasts and proposed methods. An investor can form a portfolio of six risky and one risk-free asset. The portfolio is rebalanced dynamically at the end of each period, i.e. new weights for each asset are computed. In this study we follow the methodology proposed in Della Corte, Sarno and Tsiakas (2009, 2011), Li, Tsiakas and Wang (2015) and Ahmed, Liu and Valente (2016).

More precisely, a US based investor allocates part of her wealth in the US riskless asset and the other part in the foreign risk free asset. Hence, the investor is exposed to exchange rate risk. The investor generates daily forecasts of FX returns using the predictors/ specifications presented in the previous section. Next, conditionally on the forecasts, she can solve the optimisation problem in order to rebalance the portfolio between the domestic riskless and foreign risky assets. The domestic and foreign riskless assets are the respective risk free interest rates, denoted as \( r_t^f \) and \( y_t^f \) respectively. The following optimization problem is solved:

\[
\begin{align*}
\max_{w_t} & \quad \hat{r}^p_{t+1|t} = w_t^T \hat{\mathbf{r}}_{t+1} + (1 - w_t^T \mathbf{u}) r_t^f \\
\text{subject to} & \quad (\sigma_p^*)^2 = w_t^T \Sigma_{t+1|t} w_t
\end{align*}
\]

where \( \hat{r}^p_{t+1|t} \) is the expected portfolio return, \( \hat{\mathbf{r}}_{t+1} \) is a 6x1 vector of the exchange rate expected return. The expected return of the risky asset at the end of period is equal to the return of the foreign riskless asset plus the return of the exchange rate, calculated by \( E_t[r_{t+1}] = y_t^f + \hat{r}_{t+1} \), \( \mathbf{u} \) is a 6x1 unit vector, \( \sigma_p^* \) is the target conditional volatility of the portfolio returns. Following Buncic and Piras (2016) and Li, Tsiakas and Wang (2015) we set \( \sigma_p^* = 10\% \). \( \Sigma_{t+1|t} \) is a 6x6 conditional variance-covariance matrix, \( \Sigma_{t+1|t} = (\mathbf{r}_{t+1} - \hat{\mathbf{r}}_{t+1})(\mathbf{r}_{t+1} - \hat{\mathbf{r}}_{t+1})' \). By solving the optimisation problem we get the optimal weights on the risky assets:

\[
w_t = \frac{\hat{\sigma}_p^* \Sigma_{t+1|t}^{-1} (\hat{\mathbf{r}}_{t+1} - \mathbf{u} r_t^f)}{\sqrt{C_t}}
\]

where \( \mathbf{r}_{t+1} - \mathbf{u} r_t^f \) is the 6x1 vector of excess returns and \( C_t = (\mathbf{r}_{t+1} - \mathbf{u} r_t^f) \Sigma_{t+1|t}^{-1} (\hat{\mathbf{r}}_{t+1} - \mathbf{u} r_t^f) \). Following Ahmed, Liu and Valente (2016) we winsorise the weights as \(-\mathbf{u} \leq w_t \leq 2\mathbf{u}\) in order to prevent extreme weights, while allowing the investor to take both short positions and leverage her positions.

The investor at the end of each period receives a realized return equal to

\[
r_{p,t+1} = w_t^T (\mathbf{r}_{t+1} - \mathbf{u} r_t^f) + r_t^f
\]
The gross portfolio return is $R_{p,t+1} = 1 + r_{p,t+1}$. Portfolio returns are generated by investing to all six available currencies until December 31st, 2017.

Following Della Corte, Sarno and Tsiakas (2009, 2011), Thorton and Valente (2012) and Ahmed, Liu and Valente (2016) we compute the out-of-sample performance fee that shows the investor’s gains from the proposed models over those generated by the benchmark. The average mean-variance utility function is given by:

$$U(R_p) = R_p^2 P T_{10} = R_{p,t+1}$$

where $R_p$ is the average gross portfolio return, given by $R_p = \frac{1}{T} \sum_{t=1}^{T} R_{p,t} + 1$ and $\gamma$ is the investor’s level of risk aversion. A lower $\gamma$ indicates a lower level of risk aversion. In our analysis, we examine two levels of risk aversion $\gamma = [2, 5]$. $P$ is the number of observations in the out-of-sample period. We calculate the difference between the proposed model $\hat{U}(R_p)$ and the benchmark $\hat{U}(R_{p,RW})$, which can be viewed as the Certainty Equivalent of the portfolio Returns (CER). It can also be interpreted as the maximum fee that the investor is willing to pay in order to switch from the RW portfolio to the competing model and is given by:

$$\Delta U = \hat{U}(R_p) - \hat{U}(R_{p,RW})$$

The second and most common measure of economic performance we employ is the Sharpe Ratio ($SR$) (Della Corte, Sarno and Tsiakas, 2009; DeMiguel, Garlappi and Uppal, 2009) and is given by:

$$SR = \frac{r_p - r_f}{\sigma_p}$$

where $r_p - r_f$ is the average excess return of a portfolio and $\sigma_p$ is the standard deviation of the corresponding portfolio return. We compute the annualized SR for each predictive model/specification.3 Finally, we consider the cumulative return of the portfolio at the end of the sample (end-of-period wealth). At each time step $t$ the portfolio is reinvested for a gross return $R_{p,t}$. At the end of the period the investor accumulates wealth equal to $TW = W_0 * R_{p,t+1} * R_{p,t+2} * \ldots * R_{p,P}$. For simplicity we assume that $W_0 = 1$.

5.1 Economic Evaluation Findings

It is well known that good forecasting ability does not necessarily imply profitability. We assess the economic value of each predictor and each method employing the $\Delta U$, $SR$ and $TW$ metrics.

---

3It is worth noting that SR does not take into account the effects of non-normality and it can underestimate the performance of dynamic strategies (Ahmed, Liu and Valente, 2016; Thorton and Valente, 2012; Jordeau and Rockinger, 2006; Han, 2006; Marquering and Verbeek, 2004), as it overestimates the conditional risk an investor faces at every $t$. 

15
The evaluation period is the same as the out-of-sample forecasting period, i.e. January 1st, 2004 to December 31st, 2017. The investor aims at maximizing portfolio returns with a target volatility of 10%. For robustness we assume two levels of risk aversion, $\gamma = [2, 5]$. Our results are summarised in Table 4.

A closer inspection of Table 4 reveals that positivity constraints in the forecasts play a significant role in portfolio management. $\Delta L$ and $\Delta VIX$ demonstrate significant economic gains irrespective of the level of risk aversion. We also observe poor performance for PLS, especially in the case of unconstrained forecasts. The results of POOL and PCA are similar. PCA outperforms POOL when positivity constraints are applied while POOL outperforms PCA in the unconstrained setting. By comparing the various methods, we observe that $C_{CP,+}$ and $C_{ICCP,+}$ outperform the remaining approaches. In general, increasing the level of risk aversion leads to an increase in the economic gains of $C_{CP,+}$ and $C_{ICCP,+}$. On the other hand, methods that do not use positivity constraints in general are negatively affected.

In Panel B of Table 4 the cumulative return at the end of the out-of-sample period is computed. Our results indicate that no method clearly outperforms the remaining ones although almost all predictors outperform the benchmark. It is noteworthy that $C_{CP,0}$, $C_{ICCP,0}$ and $C_{IC,+}$ in general have a poor performance since for most predictors/models the end of period cumulative returns are lower than the ones obtained by the RW benchmark. Among the individual predictors we are able to identify the very good overall performance of $\Delta L$, $MSCI$ and $\Delta VIX$. Similar to the statistical evaluation findings, POOL, PCA and PLS always provide higher cumulative returns although it seems that PCA outperforms the other two. The same qualitative results hold in terms of SR which is presented in Panel C of Table 4. Overall, the proposed $C_{ICCP,+}$ specification generates SRs that range from 0.24 (PLS) to 0.44 (PCA).

6 Robustness Checks

In this section, we evaluate the robustness of our findings. Specifically, we examine whether our results are sensitive to particular settings considered in the forecasting experiment and whether the proposed methods can provide superior forecasts when these settings change. The tests conducted are related to the control window, the data frequency and the out-of-sample period.

6.1 Monthly Frequency

In this section we examine the impact of the frequency on our results. More precisely, we change the frequency from daily to monthly observations. In this case we use a control window of 3 and 6 months. The objective is to confirm that the forecasting performance of the proposed
methods is qualitatively consistent irrespectively of the frequency. The out-of-sample period remains the same, from January 2004 and ends at December 2017.

The results are presented in Table 5. A closer inspection of Table 5 reveals that positivity constraints improve the performance of each method as in the case of the initial results. POOL, PCA and PLS show high positive $R^2_{OOS}$ values for GBP, CHF, EUR and CAD. Finally, the majority of the predictors, for all currencies except YEN, benefit in terms of $R^2_{OOS}$ values from the 6 month control window.

For the GBP the predictors with the highest $R^2_{OOS}$ values are CRB and OIL. The $R^2_{OOS}$ further increases when the PCA and PLS are considered together with positivity constraints. In general $C^+$ and $C^{CP,+}$ are the best performing specifications irrespectively of the control window used.

In the case of YEN, we observe that $C^{ICCP,+}$ outperforms the remaining methods followed by $C^{IC,+}$. Most predictors have a negative $R^2_{OOS}$ in the unconstrained setting for all methods, however, they improve when positivity constraints are applied in the cases of $C^{ICCP,+}$ and $C^{IC,+}$. Comparing the results between the 3 and 6 months control window, we observe that forecasting ability is decreased with the employment of a larger window. It is also remarkable that PCA and PLS, which were among the most robust predictors in the initial experiment, lose their forecasting ability when applied to monthly data.

Focusing on CHF, AUD and EUR, we note that there is compelling evidence that $C^{IC,+}$ and $C^{ICCP,+}$ provide the best forecasts, for all frequencies under consideration. In the case of CAD we observe that almost all $R^2_{OOS}$ are negative for $C^0$, $C^{IC,0}$, $C^{CP,0}$ $C^{ICCP,0}$. On the other hand, when positivity constraints are applied all $R^2_{OOS}$ become positive and most of them are statistically significant. We also observe that $C^{CP,+}$ significantly outperforms alternative methods in both frequencies.

Overall, the results between daily and monthly frequency data are qualitatively similar. Some predictors enhance their forecasting ability, but the methods aggregate perform the same, albeit with some remarkable high $R^2_{OOS}$ values especially for GBP.

6.2 Alternative Control Window

In this section, we examine the performance of the proposed methods after adjusting the control window to 75 days, roughly corresponding to 3 trading months. In this case, constrained predictors are much smoother than the case of the shorter window. Intuitively, at each point of time $t$, it is more difficult for predictors to exceed the maximum or minimum value of the last 75 observations than of the last 25 observations.

In our analysis, the out-of-sample period remains the same but we change the in-sample period by discarding observations in order to create the control window. The results are reported
Comparing the results of Tables 3 and 6 we observe only minor changes in the values of the $R^2_{OOS}$. For all currencies, methods that show superior forecasting ability in the initial setting continue to outperform the benchmark. Similarly, the forecasting accuracy of candidate predictors are almost unaffected by the control window.

For example, for GBP we observe a small increase in some specifications but a small decrease in POOL, PCA and PLS for most methods. Similarly, in the case of EUR, by increasing the control window we observe a small improvement in the results of $C^{CP,+}$ while we observe a small deterioration in the case of $C^{ICCP,+}$. Finally, the results of AUD confirm the difficulty in predicting the returns of the Australian Dollar.

6.3 Alternative Out–of-sample Period

In this section we examine whether a different in-sample and out-of-sample period has a significant impact on our results. Specifically, the out-of-sample period starts at January 1$^{st}$, 2009. Hence, the in-sample period is extended by 10 years and now includes the 2008-2010 financial crisis. The results are reported in Table 7. In general we observe an improvement in the performance of $C^{CP,+}$ while the $C^{ICCP,+}$ outperforms the alternative methods in most cases. Again, positivity constraints greatly improve the forecasting ability of all methods considered, while the dimensionality reduction techniques, PLS and PCA, show poor performance. On the other hand, we observe a small deterioration in the performance of each method in the cases of GPB and YEN. Furthermore, POOL rarely outperforms the benchmark although it was one of the best specifications in our initial setup.

7 Conclusions

Forecasting exchange rates on a daily frequency can be a rigorous task due to the difficulty of capturing the dynamics of such volatile series and the availability of a large number of potential predictors which are difficult to be chosen a priori. In this study we examine the forecasting ability of 14 financial predictors and three combination and dimensionality reduction techniques in the context of forecasting daily exchange rate returns of six widely traded currencies. We propose a hybrid ICCP approach and further consider positivity constraints. Our proposed methods are compared with the simple RW model, the simple linear bivariate model and the two recently developed methodologies, the IC and the CP proposed by Lin, Wu and Zhou (2017) and Pan, Pettenuzzo and Wang (2018), respectively. We also examine the impact of positivity constraints on the performance of each method.
Our results indicate that the proposed hybrid ICCP approach outperforms alternative methods in both the constrained and the unconstrained settings indicating that ICCP can be an important tool in daily FX return predictions. For all six currencies, ICCP shows higher forecasting ability in terms of $R^2_{OOS}$ and MSFE$_{adj}$. Imposing positivity constraints enhances significantly the forecasting ability of all methods. Daily CHF and AUD returns prove the most difficult to predict. Yet, in the case of $C^{ICCP,+}$ 13 and 12 predictors have a positive $R^2_{OOS}$ respectively, while this number falls to 12 and 8 respectively for the $C^{IC,+}$ and 10 and 7 respectively for $C^{CP,+}$. Finally, POOL generates consistently very good forecasts while we observe a poor performance by PLS and PCA. We also examine whether our forecasting approach delivers economic benefits and find that the economic evaluation supports the statistical results. Under the mean-variance framework we calculate three measures of economic value; the average utility, the Sharpe Ratio and the Cumulative Wealth. For robustness purposes we use two different levels of risk aversion. Our results indicate that our approach generates economically meaningful results. We observe that $C^{ICCP,+}$ and $C^{CP,+}$ outperform the remaining methods with the proposed hybrid method to deliver significant and consistent good results. Finally, we observe that PCA outperforms both POOL and PLS.

Finally, we perform a series of robustness tests including the change of the length of the control window, the frequency of the data and the out-of-sample period. Our results hold over all the robustness checks, supporting the initial findings: 1) positivity constraints in the forecasts significantly improve the forecasting ability of all predictors and combination or dimensionality reduction methods for all approaches and 2) the proposed hybrid ICCP approach can actually deliver very consistent and robust forecasts.

References


CONTACT US

Kent Business School, Sibson, University of Kent, Canterbury, Kent CT2 7FS

E: kbsinfo@kent.ac.uk
T: +44 (0)1227 827726

www.kent.ac.uk/kbs