**Factors, Filters, and Cross-Sectional Dependence:**
**The Currency Implied Volatility Surface Forecast Revisited**

Chu-Ping C. Vijverberg
College of Staten Island –CUNY and
CUNY, the Graduate Center

Wim P.M. Vijverberg
CUNY, the Graduate Center

Andrea V. Seo
Unaffiliated

12 July 2021

**Abstract**:

The extremely high correlations among the currency options in the implied volatility surface (IVS) signify a strong cross-sectional dependence. The measures to deal with strong cross-sectional dependence in a panel data set in the spatial and network literature are filters. This paper introduces new forecasting models by incorporating the filters into the implied volatility forecasting. Compared with the existing IVS forecasting models in the finance literature, these new models outperform the existing models most of the time.

**JEL Classification**: G13, G15, C58, C38

**Key words**: Implied volatility, currency options, cross-sectional dependence, forecasting

This paper has benefitted from PSC CUNY Summer Grant, # 68019-00 46.

## 1. Introduction

Since obtaining an implied volatility surface (IVS) at a given point of time is equivalent to specifying the prices of all calls and puts at that specific date, numerous models have been developed to capture the dynamics of this surface over time (Cont and Fonseca, 2002; Fengler et al., 2007; Daglish et al., 2007; Chalamandaris and Tsekrekos, 2010, 2014; Han et al., 2016). A model that captures these dynamics well presumably also yields good forecasts. In early years, researchers used linear models fitting implied volatilities (IVs) to time to maturity and moneyness in a cross-sectional setting (Dumas et al., 1998; Christoffersen and Jacobs, 2004). Their results suggest that the estimated parameters of these models are very unstable over time. Heston and Nandi (2000) estimate a nonlinear GARCH model and also find that some of the estimates are not stable. The implication is that the time variation of the implied volatilities matters for option pricing forecasts.

To model time series variation of the IVs, the literature suggests various factor models that may be organized under three different headings. First, building on the earlier work with linear cross-sectional models, Gonçalves and Guidolin (2006) develop a three-stage estimation procedure: (*i*) estimate a 5-parameter cross-sectional model of the IVS; (*ii*) treating those cross-sectional parameters as proxies of latent factors, model the time series dynamics of these factors; and (*iii*) using the forecasts of these factors as forecasts of the parameters of the IVS, generate a forecast of the future IVS. This procedure greatly improves the forecasts compared to those generated with the model of Dumas et al. (1998). Chalamandaris and Tsekrekos (2011) expand on the work of Gonçalves and Guidolin (2006) with a richer 7-parameter cross-sectional model that better captures nonlinearity in the time-to-maturity dimension of the IVS term structure. Second, following Stock and Watson (2002a), Chalamandaris and Tsekrekos (2010) generate forecasts through factor analysis. Specifically, it is a three-step process, (*i*) use all the observed IVs over all time periods to obtain factors directly by the method of principal components or another technique; (*ii*) model the time series dynamics of obtained factors; and (*iii*) by means of the estimated factor loadings obtained in the first step, turn the forecasts of these factors into a forecast of the future IVS. They find that vector autoregressive (VAR) factor models can yield a more accurate forecasts than a random walk model in the short-term, where the short-term is 1 to 5-days. Third, the diffusion index model (Stock and Watson, 2002b) consists of a two-step process: (*i*) factors are extracted from a host of variables (including historical IV values) that are candidates for explaining the dynamics of the IVS, and (*ii*) IVs are regressed on these factors in order to generate a forecasting model of the IVS. This approach has not been applied yet in the IVS forecasting literature; we will explore it in Sections 3 and 5 below.

The results of the IVS forecasts in the literature are not in line with each other. Some conclude that certain models beat the random walk models forecasts in short-term but not long-term, while others indicate that certain models beat the random walk models only in the medium/long term forecasts but not in the short-term. For example, Chalamandaris and Tsekrekos (2011) compare the 5- and 7-parameter cross-sectional factor models with the factor analytic approach. They find that none of the models used in the paper can outperform the simple random walk model for a 2-week (i.e., 10 days) prediction. However, for longer horizons, the gain in forecasting accuracy by means of the parametric models becomes significant, more so than for factor models. Exploring parametric and factor-analytic models, Chalamandaris and Tsekrekos (2014) find that forecasts generated with these models are better than those of the hard-to-beat benchmark random walk models, but only in the medium- and long-term and again more so for parametric models, which, by their nature, incorporate more information about the IVS shape. Guo et al. (2018) employ 14 models and show that these models only beat the random walk model within a forecast horizon of one week; they lose their predictive power beyond a week. Of all their models, VARC (VAR on first-differenced variables) provides better out-of-sample forecasts within a week; when the forecast horizon is beyond a week, simpler AR models do a better job. Even with these various unsettling results, most researchers agree that random walk models are hard to beat.

In this paper, we add a new perspective, born out of recent developments in spatial and network econometrics. Studies in this field have focused on the measurement and modelling of cross-sectional dependence in panel datasets where the data have structure over time and space (Chudik et al., 2011; Bailey et. al., 2016a; Bailey et. al., 2016b). In the quest for the source of dependencies in space, researchers distinguish between strong and weak cross-sectional dependence within a set of variables. If the data generating process contains a common factor and this factor cannot be averaged out, the data have strong cross-sectional dependence. This common factor may be extracted in different forms. Bailey et. al. (2016) use cross-unit averages as common factors and compare their modeling results with the principal component techniques. Essentially, a general practice in the literature is to remove the strong cross-sectional dependence by regressing the original data on either the cross-sectional means or the common factors of the data. Afterwards, researchers may deal with the weak cross-sectional dependence, if it exists, through different modeling strategies. IVS data may be set up as panel data with time to maturity and deltas, similar to the spatial or network structure over time and space. The extremely high correlations among these options with different time to maturity and deltas indicate the likely existence of strong cross-sectional dependence. So, the new ways to deal with strong cross-sectional dependence in the literature will be incorporated into our forecasting models. These cross-sectional means or common factors that cause the strong cross-sectional dependence will be referred to as filters in this paper.

The goal of this paper is to add to the literature on IVS forecasts through a different utilization of factor models in IV forecasts. We do so in the context of currency options. The common factor that generates strong cross-sectional dependence among the IVs may be captured in various ways: (*i*) with the cross-sectional mean of the IVs; (*ii*) with one or more principal component factors computed from the IVs; or (*iii*) with Morgan Stanley's Global FX Index (FXVIX), which is a weighted average of the implied volatilities of a basket of currencies and measures the sentiments in the overall currency market. The first two are common factors generated from the panel data itself; the third is an externally generated common factor.[1] Viewing a common factor as a filter allows us to separate the time series process of an IV into three parts: the process that generates the filter, the relationship between the filter and the IV, and the process that generates the filtered IV variable. The filter is common to all IVs and thus to the whole IVS, but the filter may affect different parts of the IVS differently, and the filtered IV variables may exhibit time series patterns (i.e., serial correlation) that differ from the filter and from each other. For example, we find the IVs of options with a short time to maturity to be more sensitive to the FXVIX; among short time-to-maturity options, we find that options with a lower delta put react more strongly to variations in the FXVIX (though this may be different in other time periods); and we find that the common factors (both internal and external) to exhibit less memory than the filtered IV variables.

In our analysis, we consider euro/US dollar and Canadian dollar/US dollar currency options. Many of the better-performing models in earlier research will be included as alternative models. We find that models with a filter in the form of FXVIX or the cross-sectional average forecast well, often better than models with principal component factors. We speculate that this is due to the time-varying factor loadings

[1] A precedent for the latter appears in Elhorst et al. (2020), who list business cycle effects, aggregate shocks such as oil price shocks, or changes in legislation or government policy as potential external common factors in a study of car traffic.

of principal components: relative to the fixed weights of the FXVIX or with the cross-sectional average, the time-varying factor loadings produce unwanted variation in forecasts. We also examine whether IV forecasts can be traded profitably. Both without and with trading costs, the returns to trading the IV forecasts with the common factor models beat those of the random walk model. However, profits from trading the IV forecasts are quickly eaten up by trading costs: trading may still be profitable if the cost is only 2 basis points on the implied volatility—which is unreasonably low—and profits disappear if the cost is 5 basis points—which is still quite low. We view this as being consistent with the efficient market hypothesis: information that could be acted upon is no longer profitable because of transaction costs.

The implied volatility data cover the period from 2010 to 2015. This avoids the major disruption of the financial crisis in 2007-2008 as well as the changing international environment when the Trump administration pursued significant changes in international trade relationships in 2017-2020.

In the following, Section 2 sets the scene with a few observations about euro options implied volatility. Section 3 outlines econometric models, which are divided into four groups. Section 4 describes the methods used to evaluate the forecasts of each model, using the random walk model as a baseline. Section 5 reports the empirical results, allowing for structural change by distinguishing 2010-2012 and 2013-2015. Section 6 considers the potential profits from trading implied volatility forecasts. Section 7 concludes.

## 2. Data and patterns

We use daily implied volatility (IV) data of Euro and Canadian Dollars (CAD). The data for both currencies started from January 3rd, 2010 and ended at December 31st, 2015. For each currency, we have the implied volatilities of 1-month, 2-month, 3-month, 6-month and 12-month maturities of the options for each of the following moneyness: 10-delta put, 25-delta put, ATM (at the money), 25-delta call, 10-delta call.[2] Thus, at each time point, we obtain $J = 25$ data points on the volatility surface for each currency. For simplicity, we refer to the degree of moneyness (delta) by numerical values of 10, 25, 50, 75, and 90. Thus, e.g., a delta of 10 (90) corresponds to a 10-delta put (10-delta call). This section provides a few stylized facts.

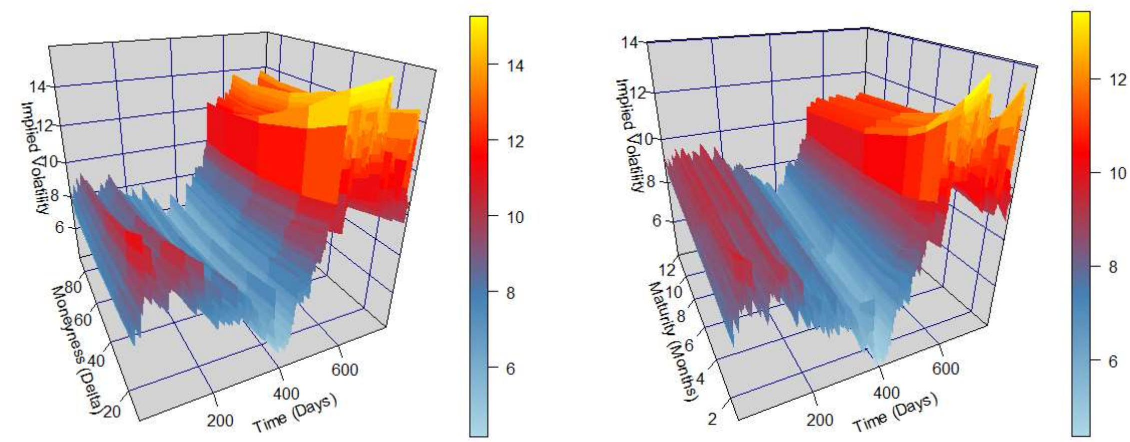
Figure 1A is the one-month maturity IV with different deltas while Figure 1B is at-the-money IV with different maturities over time. Panel A shows the typical smile, elevated on the front (low delta value) and falling toward the back (high delta value) for options at each point of time. Panel B shows that, at the earlier period, IV values do not vary much over different maturities while, at the later period, IVs with shorter maturities have higher IV values than those of longer-term maturities. The movements of IV with different maturities or different deltas have remarkably similar patterns over time. The behavior of these IV values over time is quite characteristic of a nonstationary time series. Thus, it is not surprising that all 25 time series do not pass the unit root test, i.e., all have a unit root.

[2] The Bloomberg terminal does not provide the volatilities of 10-delta put and 10-delta call directly. However, it does provide 10-delta butterfly and 10-delta Risk Reversal. The former measures [(10 delta put + 10 delta call) – ATM] while the latter measures (10-delta call – 10-delta put). With two equations and two unknowns, volatilities of 10-delta put and 10-delta call can be derived easily. The same applies to 25-delta put and call options.

Figure 1: Implied volatility of euro options over time

A: Implied volatility by delta over time (time to maturity: 1 month)

B: Implied volatility by maturity (delta: 10)




The close similarity in behavior over time is also reflected in the correlation between the time series. Table 1 presents parts of the correlation matrix, namely for three-month maturity across different delta (Panel A) and for 10-delta across different maturity. The left half of the table considers log-volatility in level; the right side considers log-volatility in first difference. The table also examines stability in these correlations by comparing 2010-2012 (below the diagonal) with 2013-2015 (above the diagonal). Correlations are as high as 0.998 and for any given maturity are nearly always above 0.90. In all this, there is little difference over time. One might expect that the correlation in first differences would be less as daily noise might take over. But correlation coefficients in the right half of the table are only slightly lower. Thus, once again, there is much evidence of a high degree of commonality in the day-to-day movements of these time series.

Methods to deal with the commonality that is evident in these time series are a part of the contribution of this paper. But how may this commonality be captured? That is what Section 3 will address.

**Table 1**: Correlation in implied volatility across options

| | Log-volatility in level | | | | | Log.volatility in first difference | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| A: Across delta for 3-month maturity options | | | | | | | | | | |
| | 10 | 25 | 50 | 75 | 90 | 10 | 25 | 50 | 75 | 90 |
| 10 | .. | 0.9986 | 0.9926 | 0.9822 | 0.9689 | .. | 0.9948 | 0.9778 | 0.9493 | 0.9152 |
| 25 | 0.9984 | .. | 0.9974 | 0.9900 | 0.9784 | 0.9889 | .. | 0.9916 | 0.9705 | 0.9403 |
| 50 | 0.9877 | 0.9942 | .. | 0.9975 | 0.9898 | 0.9684 | 0.9891 | .. | 0.9926 | 0.9726 |
| 75 | 0.9608 | 0.9721 | 0.9913 | .. | 0.9971 | 0.9321 | 0.9641 | 0.9896 | .. | 0.9904 |
| 90 | 0.9164 | 0.9308 | 0.9629 | 0.9893 | .. | 0.8878 | 0.9155 | 0.9579 | 0.9807 | .. |
| B: Across maturities for delta = 10 options | | | | | | | | | | |
| | 1 | 2 | 3 | 6 | 12 | 10 | 25 | 50 | 75 | 90 |
| 1 | .. | 0.9930 | 0.9901 | 0.9815 | 0.9606 | .. | 0.8676 | 0.8562 | 0.8133 | 0.7678 |
| 2 | 0.9943 | .. | 0.9982 | 0.9923 | 0.9741 | 0.9381 | .. | 0.9481 | 0.9163 | 0.8695 |
| 3 | 0.9862 | 0.9971 | .. | 0.9957 | 0.9799 | 0.9236 | 0.965 | .. | 0.9531 | 0.9080 |
| 6 | 0.9704 | 0.9875 | 0.9956 | .. | 0.9928 | 0.8957 | 0.9335 | 0.9571 | .. | 0.9560 |
| 12 | 0.9490 | 0.9708 | 0.9834 | 0.9939 | .. | 0.8450 | 0.8847 | 0.9118 | 0.9560 | .. |

Note: Within each 5x5 block, correlation coefficients above the diagonal pertain to 2010-2012; correlation coefficients above the diagonal pertain to 2013-2015.

## 3. Forecasting Models

Several forecasting models are examined. Some of the so-called best-performing models in the literature are selected and, at the same time, additional new models will be introduced. Let $\sigma_{it}$ denote the implied volatility of an option with delta $D$ and maturity $M$, where, depending on the context, the index $i$ is either the tuple $(D, M)$ with $D = 10, 25, 50, 75, 90$ and $M = 1, 2, 3, 6, 12$ or a simple counter $i = 1, \dots, 25$. Let $y_{it} = \ln \sigma_{it}$ be the log implied volatility. This will be the explained variable of the following analysis. In vector form, $y_t$ is a $25 \times 1$ vector stacking $y_{it}$ with $D$ running fast and $M$ running slow (unless indicated otherwise), such that options with the same maturity are grouped together.

Models are estimated on time series data running from $t = 1$ to $t = T$. Unless otherwise indicated, the disturbance $\epsilon_{it}$ has mean 0, variance $\theta_{it}^2$, and zero serial correlation. $h$-day-ahead forecasts are generated for $t = T + h$ with $h = 1,2, \dots H$. When a model is specified in first difference, the forecast $\hat{y}_{i,T+h}$ is computed as $y_{iT} + \sum_{j=1}^{h} \widehat{\Delta y}_{i,T+j}$.

The set of explored models are divided into four groups. For notation, $L$ denotes the lag operation (such that $Ly_{it} = y_{i,t-1}$), $\Delta y_{it} = (1 - L)y_{it} = y_{it} - y_{i,t-1}$, and $\Delta_h y_{it} = y_{it} - y_{i,t-h}$. Symbols are reused for different models—e.g., intercepts are always indicated as $\alpha$—and the same symbol is not meant to have the same value (or meaning or even dimension) in different models.

### *3.1. Basic Models: RW, ARMA, ARIMA*

The three models in Group I draw on standard time series techniques.

I.1 RW: Random walk

$$y_{it} = y_{i,t-1} + \epsilon_{it} \tag{3.1}$$

Forecasts are straightforward:

$$\hat{y}_{i,T+h} = \hat{y}_{i,T+h-1} = \cdots = \hat{y}_{i,T+1} = y_{it}$$

I.2 ARMA: univariate ARMA($p$,$q$) with $0 \le p \le 5$ and $0 \le q \le 3$.

$$\Phi_i(L)y_{it} = \alpha_i + \Theta_i(L)\epsilon_{it} \tag{3.2}$$

where $\Phi_i(L)$ and $\Theta_i(L)$ are polynomials of order $p$ and $q$, respectively. The choice of parameters is optimized for each window by means of the BIC criterion. The maximum of $p$ is motivated by the fact that data are measured by business day: a week represents five consecutive observations. For the purpose of forecasting, we may rewrite this model as

$$y_{it} = \alpha_i + \phi_{1i}y_{i,t-1} + \cdots + \phi_p y_{i,t-p} + \epsilon_{it} + \theta_1\epsilon_{i,t-1} + \cdots + \theta_q\epsilon_{i,t-q}$$

Forecasting is accomplished by rolling this equation forward.

I.3 ARIMA: univariate ARIMA($p$,1,$q$) with $0 \le p \le 5$ and $0 \le q \le 3$.

$$A_i(L)(1-L)y_{it} = \alpha_i + B_i(L)\epsilon_{it} \tag{3.3}$$

The choice of parameters $p$ and $q$ is optimized for each window by means of the BIC criterion. This model is of course equivalent to an ARMA($p$,$q$) model of $\Delta y_{it}$. Forecasting of $y_{i,T+h}$ is accomplished by rolling this equation forward to obtain $\widehat{\Delta y}_{i,T+h}$ and accumulating the computed first differences.

Vector autoregression (VAR) models of various kinds have been proposed in the literature as well; for example, see Chalamandaris and Tsekrekos (2010) and Guo et al. (2017). Since they are highly parameter-intensive and therefore underperform in comparison with the random walk approach, they are omitted from the list of standard time series model in this section. See Appendix A.2 for a description of them and their performance.

### *3.2. Day-of-Week Seasonality Models: AR*(5), ARI$^S$*

Group II consists of two models that focus on the fact that option contracts are traded daily on weekdays.

II.1 AR*(5) and AR*(5)D: Day of week effects

The data on option volatility are gathered daily on business days. This creates the potential for a day-of-week pattern in the behavior of the outcome variable. On Fridays, market participants prepare their positions both to close the week and to bridge the weekend; on Mondays, market participants respond to the news that accumulated over the weekend. Thus, the model is stated as:[3]

$$\Delta y_{it} = \alpha_i + D_t^s\gamma_i + \epsilon_{it} \tag{3.4}$$

where $D_t^s$ is a vector of four dummy variables to denote Monday through Thursday, with Friday as the base day. $\epsilon_{it}$ follows an AR process, with the lag length acknowledging the length of the business week:

$$\epsilon_{it} = \phi_{1i}\epsilon_{i,t-1} + \cdots + \phi_{5i}\epsilon_{i,t-5} + \nu_{it} \tag{3.5}$$

Explorations with the data showed that $\hat{\phi}_{3i}$ and $\hat{\phi}_{4i}$ were frequently statistically insignificant whereas $\hat{\phi}_{5i}$ does contribute for about half of the 25 equations. For reason of parsimony, $\phi_{3i}$ and $\phi_{4i}$ are therefore a

[3] A similar model might be formulated in levels. However, finance theory would argue against a regular day of week effect in the level of volatility. Empirically, $\hat{\gamma}_i$ never gains statistical significance during the time period under study in a model where the dependent variable is $y_{it}$.

priori set to 0, yielding what will be referred to as an AR*(5) model (if omitting $D_t^s$ from equation (3.4)) and an AR*(5)D model (if including $D_t^s$).

II.2: ARI$^S$($p$,$d$,$P_s$): Weekly-seasonal AR

The AR*(5) model states that option volatility tends to move up more (or down less) on certain days of the week than on other days, and that the magnitude of this weekly pattern remains constant over time. A multiplicative seasonal model is usually expressed as ARIMA($p$,$d$,$q$)×($P_s$,$D_s$,$Q_s$). To keep the model parsimonious, we consider only a seasonal component of order 1 and an AR component of order 2, at most. Thus, an ARI$^S$($p$,$d$,$P_s$) model is introduced. For example, for $p = 2$, $d = 1$, $P_s = 1$, we have

$$\Delta y_{it} = \alpha_i + \epsilon_{it} \tag{3.6}$$

$$(1 - \xi_i L^5)(1 - \phi_{1i} L - \phi_{2i} L^2)\epsilon_{it} = \nu_{it} \tag{3.7}$$

### *3.3. Factor Models: CFLk, CFCk, GG, DIk, DIkF and DIkFD*

Group III consists of six models suggested in the literature, all based on factor analysis.

III.1 CFLk: Common factor in levels with $k$ factors

As illustrated in Section 2, the time series behavior of the $y_{it}$ is highly correlated. Following in the footsteps of Bai and Ng (2002) and Pesaran (2015), it may be beneficial to extract one or more common factors $F_t = (F_{1t}, \dots, F_{kt})$ from the 25 time series $(y_{1t}, \dots, y_{25,t})$, decompose each $y_{it}$ into a part that is a function of these common factors and a remainder. Specifically, let $E$ be the $(25 \times k)$ matrix of factor loadings (equal to the $k$ eigenvectors belonging to the $k$ largest eigenvalues of the correlation matrix of $y_t$), and let $E_i$ be the $i$th row of $E$. Let $\mu_i$ and $\sigma_i$ be the sample mean and standard deviation of $y_{it}$. Then

$$y_{it} = \mu_i + \sigma_i F_t E_i' + \epsilon_{it} \tag{3.8}$$

and therefore, after $\hat{F}_{T+h}$ is forecast with an optimized ARIMA($p$,$d$,$q$) model:

$$\hat{y}_{i,T+h} = \mu_i + \sigma_i \hat{F}_{T+h} E_i' \tag{3.9}$$

III.2 CFCk: Common factor in first difference

As $y_{it}$ is highly correlated with each other, so is $\Delta y_{it}$. The CFC model parallels the CFL model by formulating it terms of $\Delta y_{it}$:

$$\Delta y_{it} = \mu_i + \sigma_i F_t E_i' + \epsilon_{it} \tag{3.10}$$

where $\mu_i$ and $\sigma_i$ now are the sample mean and standard deviation of $\Delta y_{it}$.

This first-difference model may be extended with the day of week dummies,[4] where $F_t$ is computed from the correlation matrix of the day-of-week-filtered $\Delta y_{it}$:

$$\Delta y_{it} = \alpha_i + D_t^s \gamma_i + \sigma_i F_t E_i' + \epsilon_{it} \tag{3.11}$$

This model will be referred to as CFCkD.

III.3 GG with optimized ARIMA($p$,$d$,$q$)

[4] In principle, common factors $F_t$ combined with factor loadings $E_i'$ in the CFC model could capture day-of-week effects that are modeled with the $D_t^s \gamma_i$ in equation (3.11). The advantage of modeling the day-of-week effects explicitly is that $F_t$ is data-driven and therefore varies between data windows with the shocks occurring during those windows, whereas $D_t$ is deterministic and therefore parsimonious.

Gonçalves and Guidolin (2006) modeled the implied volatility surface in terms of moneyness and time to maturity at each point in time:

$$y_{it} = \beta_{0t} + \beta_{1t} D_i + \beta_{2t} D_i^2 + \beta_{3t} M_i + \beta_{4t} D_i M_i + \epsilon_{it} \equiv x_i' \beta_t + \epsilon_{it} \tag{3.12}$$

This model is estimated with OLS, which in principle assumes that $\epsilon_{it}$ is distributed iid$(0, \sigma_{it}^2)$. It would seem that $\hat{y}_{i,T+h}$ could be computed as $x_i' \hat{\beta}_{T+h}$, inserting $\hat{\epsilon}_{i,T+h} = 0$. The latter turns out to be a poor strategy. For some points $(M_i, D_i)$ on the implied volatility surface, the distribution of residuals $\hat{\epsilon}_{it}$ or, more illustrative, $\hat{\epsilon}_{DMt}$ over time does not center at 0. This may be viewed as a misspecification of equation (3.12), but the effect of this misspecification can be forecasted. Thus, forecasts are generated as

$$\hat{y}_{i,T+h} = x_i' \hat{\beta}_{T+h} + \hat{\epsilon}_{i,T+h} \tag{3.13}$$

and each element of $\hat{\beta}_{T+h}$ and each $\hat{\epsilon}_{i,T+h}$ is forecasted with an optimized univariate ARIMA($p$,$d$,$q$) model. In preliminary analysis, $\beta_{0t}$ was to have a unit root; the other slopes were stationary; and residuals never have a unit root either. Therefore, $\beta_{0,T+h}$ is forecasted with an optimized ARIMA($p$,1,$q$) model and slopes $\beta_{j,T+h}$ for $j = 1, \ldots, 4$ and $\epsilon_{it}$ are forecasted with an optimized ARIMA($p$,0,$q$) model, in both cases with $0 \leq p \leq 5$ and $0 \leq q \leq 3$.

The GG model is in effect a factor model: the values of $y_{it}$ for $i = 1, \ldots, 25$ are condensed into five $\beta_{jt}$ at each $t$, with preselected factor loadings that were named $x_i$ in the discussion above. More specifically, apart from scaling, $x_i$ in equation (3.9) fulfills the same function as $E_i$ in equation (3.12). However, a simple forecast $\hat{y}_{i,T+h} = x_i' \hat{\beta}_{T+h}$ is not competitive as $\hat{\epsilon}_{it}$ still carries information that can be used for forecasting; this parallels the presence of $\mu_i$ in equation (3.9).

### III.4 Diffusion index models: DIk, DIkF and DIkFD

Stock and Watson (2002ab) introduced a diffusion index model for forecasting. The model is related to factor models, which can be summarized as the following. Let $Z_t$ be a $(1 \times l)$ row vector of stationary variables relevant for the forecasting of $y_{it}$. The variables are condensed into $k$ factors represented by the row vector $F_t$, with $E$ as the $m \times k$ matrix of factor loadings:

$$Z_t = F_t E' + \varepsilon_{zt} \tag{3.14}$$

Once $F_t$ is computed, estimate an ARMAX(1,0) model that relates each $\Delta y_{it}$ to the diffusion indices:

$$\Delta y_{it} = \alpha_i + F_t \beta_i + \phi_i \Delta y_{i,t-1} + \epsilon_{it} \tag{3.15}$$

We implement $Z_t$ in three different forms. In the first, $Z_t$ includes the history of all the first-differenced implied volatilities $\Delta y_t$, taken to consist of its first five lags: $Z_t = (\Delta y_{t-1}', \ldots, \Delta y_{t-5}')$ such that $l = 125$. The second form augments this with the history of other financial variables $\Delta f_t$, again captured by their first five lags: $Z_t = (\Delta y_{t-1}', \ldots, \Delta y_{t-5}', \Delta f_{t-1}', \ldots, \Delta f_{t-5}')$. The third form adds the day of the week the option is traded through the row vector $D_t^s$ defined in equation (3.4) (four dummy variables to denote Monday through Thursday, with Friday as the base day): $Z_t = \{\Delta y_{t-1}', \ldots, \Delta y_{t-5}', \Delta f_{t-1}', \ldots, \Delta f_{t-5}', D_t^s\}$. Thus, letting $k$ be the number of factors retained, the three models under the diffusion index rubric will be referred to as DI$k$ if $Z_t$ contains only past first-differences in $y$; as DI$k$F if $Z_t$ adds financial variables; and as DI$k$FD if if $Z_t$ also contains a day-of-week component.

Consider forecasting with the DI$k$ model. To forecast $\Delta y_{i,T+h}$, a forecast of $F_{T+h}$ is needed. First, note that $F_{T+1}$ is derived from $Z_{T+1} = (\Delta y_T', \ldots, \Delta y_{T-4}')$, which is therefore fully observed. Therefore,

$$\widehat{\Delta y}_{i,T+1} = \hat{\alpha}_i + F_{T+1}\hat{\beta}_i + \hat{\phi}_i \Delta y_{iT} \tag{3.16}$$

To forecast $F_{T+2}$, we make use of the just-obtained forecast of $\Delta y_{T+1}$. Thus, $\hat{z}_{T+2} = \left(\widehat{\Delta y}_{T+1}, \Delta y_T, \dots, \Delta y_{T-3}\right)$ and therefore also $\hat{F}_{T+2} = \hat{F}(\hat{z}_{T+2})$. It follows that

$$\widehat{\Delta y}_{i,T+2} = \hat{\alpha}_i + \hat{F}_{T+2}\hat{\beta}_i + \hat{\phi}_i \widehat{\Delta y}_{i,T+1} \tag{3.17}$$

This process of rolling forward continues until the end of the forecast horizon.

In the case of DI$k$F, $z_t = (\Delta y_{t-1}, \dots, \Delta y_{t-5}, \Delta f_{t-1}, \dots, \Delta f_{t-5})$. The forecast for period $T+1$ is as above, but forecasts cannot be simply rolled forward since $\Delta f_{T+1}$ is unknown and must be separately forecasted). Thus, forecasting necessitates an optimized ARMA($p$,$q$) forecasting step to obtain each $F_{T+h}$.

In the case of DI$k$FD, the day of week enters in two places. First, note that in its common form, factor analysis starts with standardizing $z_t$. Each element of a vector in $z_t$ is taken in deviation from the same mean. Now, to account for the day of week, take each element of a vector in $z_t$ in deviation of the trading-day-specific mean. In essence, this implies that $F_t = \hat{F}(z_t, D_t^s)$. Second, add $D_t^s$ to the relationship of $y_{it}$ with $F_t$:

$$\Delta y_{it} = \alpha_i + D_t^s \gamma_i + F_t \beta_i + \phi_i \Delta y_{i,t-1} + \epsilon_{it} \tag{3.18}$$

$\gamma_i$ is a $(4 \times 1)$ vector of slopes of day-of-week intercept shifters. The remainder of the steps are similar to the process followed for DI$k$F.

### *3.4. Filter-Related Models: AVEL, AVEC, FXVIXL, FXVIXC*

Apart from $\beta_t$ in the GG model and $F_t$ in the CFL and CFC models, what other common factor models could be formulated? Four come to mind, which are inspired by the familiar CAPM model. We will refer to these factors as "observed common factors." The models constitute Group IV.

As shown in Section 2, the correlations among the 25 options are extremely high. These high correlations exist not just for the levels, but also for the first-differenced of these options. These correlations illustrate the strong cross-sectional dependence in the IVS of FX options. The common practice in the recent literature (Bailey et. al., 2016; Elhorst et al., 2020) is to remove the strong cross-sectional dependence by regressing the original data on either the cross-sectional means or the common factors of the data or on another observed variable. These variables will be referred to as filters in this paper. The "filter" concept will be incorporated in the following forecasting models.

IV.1 AVEL: Average Implied Volatility, in level

The first is extremely simple, yet quite powerful for forecasting relative to the CFL model: the cross-sectional average of the options' implied volatilities, or $\bar{y}_t = \frac{1}{25}\sum_{i=1}^{25} y_{it}$. For each $i$, we formulate

$$y_{it} = \alpha_i + \delta_i \bar{y}_t + \tilde{y}_{it} \tag{3.19}$$

$\tilde{y}_{it}$ may be thought of as the implied volatility of option $i$ filtered by the average implied volatility across all options under consideration. Estimates of $\delta_i$ that are much greater than $\frac{1}{25}$ are evidence of the high degree of commonality among the $y_{it}$.[5] Forecasting of $y_{i,T+h}$ necessitates generating a forecast of $\bar{y}_{T+h}$ and a forecast of $\tilde{y}_{i,T+h}$.

[5] Estimates of $\alpha_i$ and $\delta_i$ suffer from endogeneity bias, but this is of no concern when the objective is to obtain accurate forecasts: the forecasting equation suffers from the same bias as the estimation equation.

$$\hat{y}_{i,T+h} = \hat{\alpha}_i + \hat{\delta}_i \hat{\bar{y}}_{T+h} + \hat{\tilde{y}}_{i,T+h} \tag{3.20}$$

Why might this be beneficial relative to the approaches in Groups I and III? Here, the time series process of the common factor is separated from the time series process of each specific option. Therefore, a single time series model does not have to represent two separate processes. Thus, the processes for $\bar{y}_t$ and $\tilde{y}_{it}$ may have different lag lengths and may depend in different ways on the day of week ($D_t$). In the implementation of this model, the prediction model for $\bar{y}_t$ is generated with an optimized ARIMA($p$,$d$,$q$) model with $1 \leq p \leq 5$, $d = \{0{,}1\}$, and $0 \leq q \leq 3$, For the process for $\tilde{y}_{it}$, two models are considered as the filtering process removes the unit root from some but not all $y_{it}$: AVEL1 specifies an optimized ARIMA model, and AVEL2 specifies an optimized ARI$^S$($p$,$d$,$P$) model with $1 \leq p \leq 2$, $d = \{0{,}1\}$ and $P \leq 1$.

IV.2 AVEC

Formulated in first difference terms, the AVEC model blends the AR(2) model with a day-of-week element and the common factor that comes from the change in the average option volatility.

$$\Delta y_{it} = \alpha_i + D_t^S \gamma_i + \delta_i \Delta \bar{y}_t + \phi_{1i} \Delta y_{i,t-1} + \phi_{2i} \Delta y_{i,t-2} + \epsilon_{it} \tag{3.21}$$

This model is closely related to the AVEL model, of course, but allows a constant day-of-week pattern rather than a seasonal AR pattern that allows from dissipating day-of-week effects. The $\Delta\bar{y}_t$ process is estimated separately as an AR*(5) (an AR(2) model augmented with a fifth lag) with day-of-week:

$$\Delta \bar{y}_t = \alpha_i^a + D_t^S \gamma_i^a + \phi_{1i}^a \Delta \bar{y}_{t-1} + \phi_{2i}^a \Delta \bar{y}_{t-2} + \phi_{5i}^a \Delta \bar{y}_{t-5} + \epsilon_{it}^a \tag{3.22}$$

IV.3 FXVIXL

The third CAPM-inspired model broadens the common factor to the overall volatility measure FXVIX in currency markets, shown as $f_{xt}$ in the following equation:

$$y_{it} = \alpha_i + \delta_i f_{xt} + \tilde{y}_{it} \tag{3.23}$$

where $\tilde{y}_{it}$ is now the implied volatility of option $i$ filtered by the FXVIX of the entire market for currencies. In the first implementation (FXVIXL1), the prediction model for $f_{xt}$ is generated with an ARIMA(1,1,0) with day-of-week effects; $\tilde{y}_{it}$ is estimated with an optimized ARI$^S$($p$,$d$,$P$) model with $1 \leq p \leq 2$, $d = \{0{,}1\}$ and $P \leq 1$. The filtering with $f_{xt}$ does not always remove the unit root from $y_{it}$. In the second implementation (FXVIXL2), optimized ARIMA($p$,$d$,$q$) processes are estimated for both $f_{xt}$ and $\tilde{y}_{it}$.

IV.4 FXVIXC

The FXVIXC model parallels the AVEC model: it blends the AR($p$) model with a day-of-week element and the contextual factor that comes from the change in the FXVIX:

$$\Delta y_{it} = \alpha_i + D_t \gamma_i + \delta_i \Delta f_{xt} + \phi_{1i} \Delta y_{i,t-1} + \phi_{2i} \Delta y_{i,t-2} + \phi_{5i} \Delta y_{i,t-5} + \epsilon_{it} \tag{3.24}$$

The third and fourth lags of an AR(5) model are rarely statistically significant, but the fifth lag is, thus giving rise to the AR*(5) serial correlation structure. $\Delta f_{xt}$ is estimated with an AR(1) process augmented with a day-of-week element.

## 4. Out-of-Sample Forecast Evaluation Methods

As mentioned at the start of Section 3, models are estimated on time series data running from $t = 1$ to $t = T$, and $h$-day-ahead forecasts are generated for $t = T + h$ with $h = 1, 2, \ldots, H$. For comparison purposes, let rolling time windows be indexed by $n = 1, \ldots, N$, and let $\tilde{t}$ denote calendar time.

Thus, window $n$ utilizes data from $\tilde{t} = (n-1)+1$ through $\tilde{t} = (n-1)+T$, generating forecasts for $\tilde{t} = (n-1)+T+1$ through $\tilde{t} = (n-1)+T+H$. In this way, $h$-period-ahead forecasts of different models may be evaluated on the basis of $N$ forecast attempts. In tabulations in the following sections, we use $T = 504$ (two years of daily data) and $N = 252$ (one year's worth of $h$-period-ahead forecasts). In the following, every forecast and forecast error ought to be subscripted with $n$, but this subscript is omitted for notational clarity.

Unless indicated otherwise, the baseline model is the random walk. Let $e_{i,T+h}^{rw}$ be the $h$-period-ahead forecast error of the random walk model, and let $e_{i,T+h}$ be the same forecast of a comparison model:

$$e_{i,T+h}^{rw} = y_{i,T+h} - \hat{y}_{i,T+h}^{rw} \tag{3.25}$$

$$e_{i,T+h} = y_{i,T+h} - \hat{y}_{i,T+h} \tag{3.26}$$

Again, note that these are computed for each rolling window $n$ with $n = 1, \ldots, N$. $e_{i,T+h}$ and $e_{i,T+h}^{rw}$ are the key ingredients of the out-of-sample $R^2$, the Diebold-Mariano and Clark-West tests, and the Sign test that are discussed in the next subsections.

### *4.1. Out-of-Sample $R^2$*

The primary descriptive statistic is the out-of-sample $R_{ih}^2$, defined for a given forecasting model for each option $i$ and forecast horizon $h$:

$$R_{ih}^2 = 1 - \frac{\sum_{n=1}^{N} e_{i,T+h}^2}{\sum_{n=1}^{N} \left(e_{i,T+h}^{rw}\right)^2} \tag{3.27}$$

It measures the proportional improvement in forecasts by a model (say, AVEC) at horizon $h$ relative to the random walk model. A positive value indicates that the AVEC model generates smaller forecast errors and thus is more accurate; a negative value shows that the random walk model is preferable.

To test whether this improvement is statistically significant, it matters whether the random walk model is nested within a given model. For example, the ARMA(p,q) model reduces to the random walk model if, in equation (3.2), $\phi_{1i} = 1$ and all other parameters equal 0; the VARL model reduces to the random walk model if $\alpha = 0$ and $A = I_{25}$; the GG model cannot be reduced to the random walk model by any parameter restriction. If a model does not nest the random walk model, the modified Diebold-Mariano (MDM) test applies (Diebold and Mariano, 1995; West, 1996; Harvey et al., 1997; Rapach and Wohar, 2007); if a model does nest the random walk model, a modified Clark-West (MCW) test applies (Clark and West, 2007, Harvey et al., 1997).

### *4.2. The Modified Diebold-Mariano (MDM) and Clark-West (MCW) tests*

Suppose that the comparison model does not nest the random walk model. Define

$$g_{ihn} = \left(e_{i,T+h}^{rw}\right)^2 - e_{i,T+h}^2 \tag{3.28}$$

Let $\bar{g}_{ih} = \frac{1}{N}\sum_{n=1}^{N} g_{ihn}$. Let $\mu_{g_{ih}}$ denote the population mean of $g_{ihn}$. The Null hypothesis states that the two models have equal predictive ability: $H_0\colon \mu_{g_{ih}} = 0$. Under the Null hypothesis, we have by Diebold and Mariano (1995) and Clark (1997):

$$MDM_{ih} = \frac{N+1-2h+h(h-1)/N}{N} \times \frac{\bar{g}_{ih}}{V(\bar{g}_{ih})^{1/2}} \sim t(N-1) \tag{3.29}$$

where $V(\bar{g}_{ih}) = \frac{1}{N}\left(V_{0h} + \sum_{j=1}^{h-1} V_{jh}\right)$ and $V_{jh} = \frac{1}{N}\sum_{n=j}^{N}(g_{ihn} - \bar{g}_{ih})(g_{ih,n-j} - \bar{g}_{ih})$ for $j = 0, \ldots, h-1$. The alternative hypothesis is two-sided ($H_A: \mu_{g_{ih}} \neq 0$): one is better than the other. Thus, a significant positive value provides evidence that the alternative model yields more accurate forecasts; a significant negative value indicates that the random walk model is preferable.

If the model does nest the random walk model, $g_{ihn}$ is modified to account for estimation error that adds to the forecast error of the random walk model (Clark and West, 2007):

$$g_{ihn} = \left(e_{i,T+h}^{rw}\right)^2 - \left(e_{i,T+h}^2 - \left(e_{i,T+h}^{rw} - e_{i,T+h}\right)^2\right) = e_{i,T+h}^{rw}\left(e_{i,T+h}^{rw} - e_{i,T+h}\right) \tag{3.30}$$

Clark and McCracken (2013) recommend the following test statistic, which they call ENC-t and CW-t, modified again as suggested by Harvey et al. (1997) and thus called MCW:

$$MCW_{ih} = \frac{N+1-2h+h(h-1)/N}{N} \times \frac{\bar{g}_{ih}}{V(\bar{g}_{ih})^{1/2}} \sim t(N-1) \tag{3.31}$$

The Null hypothesis states that the two models have equal predictive ability in population: $H_0: \mu_{g_{ih}} = 0$. Thus, the alternative is now one-sided ($H_A: \mu_{g_{ih}} > 0$) since, when the true population parameters are inserted, the alternative model may be able to generate better forecasts than the random walk model, but the random walk model can never generate better forecasts than the alternative model.

### *4.3 Out-of-sample $R^2$ and MDM/MCW tests in grouped settings*

The tools in the previous subsections are also applicable in group settings. Let $g_{hn}$ be the $25 \times 1$ vector stacking $g_{ihn}$, and let $\iota_{25}$ be the vector (with elements equal to 0 and 1) that defines the group. Define $\tilde{g}_{hn} = \iota_{25}' g_{hn}$ and $\bar{\tilde{g}}_h = \frac{1}{N}\sum_{n=1}^{N} \tilde{g}_{hn}$. Then, the test is implemented along the same lines as equations (3.29) and (3.31). The Null hypothesis here is $H_0: \mu_{\tilde{g}_h} = 0$. The meaning of this hypothesis depends on whether the alternative model nests the random walk model. In the nested case, $H_0: \mu_{\tilde{g}_h} = 0$ is equivalent to $H_0: \mu_{g_{ih}} = 0$ for all $i$ in the group. Thus, it constitutes a joint test of whether the alternative model outperforms the random walk model for any option in the group. For the non-nested case, it is possible that the $H_0: \mu_{\tilde{g}_h} = 0$ is true but $\mu_{g_{hi}} > 0$ for one option in the group and $\mu_{g_{hi}} < 0$ for another option.

In this regard, recall that the out-of-sample $R_{ih}^2$ measures the relative improvement in forecast performance of the alternative model for option $i$ at horizon $h$. A grouped out-of-sample R2-type measure may be written as

$$\tilde{R}_h^2 = 1 - \frac{\sum_{i=1}^{25} \iota_{25,i} \sum_{n=1}^{N} e_{i,T+h}^2}{\sum_{i=1}^{25} \iota_{25,i} \sum_{n=1}^{N} \left(e_{i,T+h}^{rw}\right)^2} = 1 - \sum_{i=1}^{25} w_i \frac{\sum_{n=1}^{N} e_{i,T+h}^2}{\sum_{n=1}^{N} \left(e_{i,T+h}^{rw}\right)^2} = \sum_{i=1}^{25} w_i R_{ih}^2 \tag{3.33}$$

where $w_i = \sum_{n=1}^{N}\left(e_{i,T+h}^{rw}\right)^2 / \sum_{i=1}^{25} \iota_{25,i} \sum_{n=1}^{N}\left(e_{i,T+h}^{rw}\right)^2$ is the share of the sum of squared forecast errors of option $i$ as a fraction of the sum of squared forecast errors of the group to which it belongs. Thus, $\tilde{R}_h^2$ is a weighted average of the $R_{ih}^2$ of the currency options in the group. In Section 5, we employ this grouped $\tilde{R}_h^2$ to evaluate the forecasts of all options with forecast horizon $h$ together: in this case, $\iota_{25,i} = 1$ for all $i$.

### *4.4. Sign test*

A second test to distinguish between forecasting models is the sign test (Diebold and Mariano, 1995; Kitchens, 2002; Gibbons and Chakraborti, 2003).[6] Define the indicator function $I(g_{ihn}) = 1$ if $g_{ihn} \geq 0$ and $= 0$ if $g_{ihn} < 0$. The sign test may then be written as:

$$ST_{ih} = \frac{2}{N^{0.5}}\left(\sum_{n=1}^{N} I(d_{ihn}) - 0.5N\right) \sim^{a} N(0,1) \tag{3.34}$$

A significant negative value is evidence in favor of the alternative model; a significant positive value indicates that the random walk model is preferable.

## 5. Forecasting performance

### *5.1. An overall assessment*

In this section, each model's forecasting performance will be evaluated by grouping all 25 option IVs together. Table 2 provides the evaluations by means of five summary statistics: (*i*) $\tilde{R}_h^2$ for each forecast horizon; (*ii*) significance levels of MDM/MCW tests as indicated by "*" and "+" markers: one-sided tests in favor of indicated model are marked as *** (1%), ** (5%), and * (10%), whereas one-sided tests in favor of random walk model are shown as +++ (1%), ++ (5%), and + (10%); (*iii*) the percentage of the MDM/MCW tests of individual option volatility models that are statistically significantly different from the random walk at the 10% level, across all horizons, shown under the "Share preferred" heading; (*iv*) for those models that do not nest the random walk model,[7] the percentage of the MDM tests of individual option volatility models that indicate superior performance of the random walk at the 10% significance level, across all horizons; and (*v*) the percentage of the sign tests of individual option volatility models that are statistically significantly different from the random walk at the 10% level, across all horizons, shown under the heading of "Sign tests."

To take the second line of the table as an example, when the implied volatility models are estimated with an optimized ARIMA model, the overall out-of-sample $R_h^2$ at a forecast horizon of $h = 1$ equals 0.008, indicating that ARIMA's mean squared prediction error is 0.8% smaller than that of the random walk model. For ARIMA, the null hypothesis is that MSE of ARIMA is greater than or equal to MSE of random walk. The *** marker indicates rejection of the null at the 1% significance level. Thus, in the aggregate, ARIMA is a better forecast model than random walk. At longer forecast horizons ($h \geq 2$), the mean squared prediction error of ARIMA tends to be slightly larger. Summarizing IV-level tests, as seen under the "Share preferred" heading, the left column ("Model") shows that 28.57 percent of the 175 (25 option volatility series at seven forecast horizons) ARIMA forecast performances are better at the 10 percent significance level than random walk. The entry in the right column ("R.Walk") is empty ("n.a.") because the random walk model is nested within the ARIMA model. The sign test reveals that in 8.57 percent of the 175 comparisons the ARIMA model performed better than the random walk model at the 10 percent significance level.

The second group of models examine day-of-week effects. The AR*(5) model is a baseline specification for these four rows. It models $\Delta y_{it}$ with an autoregressive model that contains the first,

---

[6] The Wilcoxon Signed Rank Test has more power than the Sign Test but requires symmetry in $d_{ihn}$. Symmetry is too often rejected for some models, for some options $i$, and/or for some forecast windows $h$.

[7] The factor-based models (CFCk, CFLk, GG, and DIk and their variants) do not nest the random walk specification as there is no parameter restriction that would reduce the model to the random walk formulation. All other models contain the random walk specification as a special case.

**Table 2**: Aggregate out-of-sample R2 and MDM/MCW tests, and summary of separate MDM/MCW and sign tests (euro, 2010-2012)

| Type of Model | $\tilde{R}_h^2$ by forecast horizon[a] $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Share preferred[b] Model | R.Walk | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| I. Basic time series models | | | | | | | | | | |
| ARMA | -0.023 | -0.064 | -0.162 | -0.309 | -0.401 | -0.484 | -0.536 | 4.57 | n.a. | 0.00 |
| ARIMA | 0.008 *** | -0.007 | -0.005 | -0.006 | -0.002 | -0.001 | 0.000 | 28.57 | n.a. | 8.57 |
| II. Day-of-week seasonality models | | | | | | | | | | |
| AR*(5) | 0.016 *** | 0.003 ** | -0.004 | -0.011 | -0.020 | -0.026 | -0.034 | 14.86 | n.a. | 23.43 |
| AR*(5)D | 0.056 *** | 0.040 *** | -0.005 | -0.014 | -0.021 | -0.026 | -0.032 | 22.86 | n.a. | 26.29 |
| ARI$^S$($p$,$d$,$P_s$) | 0.014 *** | 0.002 * | -0.005 | -0.013 | -0.021 | -0.025 | -0.033 | 16.00 | n.a. | 25.71 |
| ARI$^S$(2,1,1) | 0.016 *** | 0.003 ** | -0.005 | -0.013 | -0.022 | -0.027 | -0.037 | 15.43 | n.a. | 20.00 |
| III. Factor models | | | | | | | | | | |
| CFC3 | 0.010 | -0.012 | 0.002 | -0.006 | -0.015 | -0.021 | -0.031 | 0.57 | 2.29 | 33.14 |
| CFC3D | 0.062 ** | 0.036 * | 0.003 | -0.006 | -0.014 | -0.021 | -0.030 | 6.86 | 0.00 | 34.29 |
| CFL3 | -0.229 +++ | -0.106 +++ | -0.034 ++ | -0.006 | -0.001 | 0.003 | 0.003 | 13.71 | 29.71 | 23.43 |
| CFL10 | 0.028 * | 0.009 | 0.014 | 0.017 | 0.013 | 0.014 | 0.011 | 28.00 | 7.43 | 31.43 |
| GG | -0.025 | -0.038 | -0.055 | -0.050 | -0.032 | -0.022 | -0.018 | 4.57 | 15.43 | 32.00 |
| DI5 | 0.009 | -0.021 | -0.017 | -0.020 | -0.026 | -0.031 | -0.037 | 0.00 | 1.71 | 15.43 |
| DI10 | 0.013 | -0.033 | -0.054 | -0.041 | -0.034 | -0.039 | -0.043 | 0.00 | 6.29 | 10.86 |
| DI5F | 0.011 | -0.034 + | -0.041 ++ | -0.023 | -0.037 | -0.040 | -0.043 | 0.00 | 16.00 | 13.71 |
| DI10F | 0.024 | -0.035 + | -0.052 +++ | -0.031 | -0.044 | -0.044 | -0.045 | 1.14 | 17.14 | 13.14 |
| DI1FD | 0.062 ** | 0.036 * | 0.004 | -0.005 | -0.014 | -0.021 | -0.030 | 7.43 | 0.00 | 30.86 |
| DI5FD | 0.058 * | 0.020 | -0.047 ++ | -0.034 | -0.036 | -0.041 | -0.041 | 6.29 | 10.86 | 26.29 |
| DI10FD | 0.066 ** | 0.021 | -0.054 ++ | -0.038 | -0.039 | -0.041 | -0.041 | 5.71 | 12.00 | 29.14 |
| IV. Filter-related models | | | | | | | | | | |
| FXVIXC | 0.056 *** | 0.041 *** | -0.002 | -0.008 | -0.016 | -0.021 | -0.026 | 25.14 | n.a. | 39.43 |
| AVEC | 0.068 *** | 0.043 *** | -0.003 | -0.013 | -0.021 | -0.026 | -0.033 | 26.29 | n.a. | 25.71 |
| FXVIXL1 | 0.017 *** | 0.033 *** | 0.061 *** | 0.113 *** | 0.139 *** | 0.164 ** | 0.190 ** | 72.57 | n.a. | 72.57 |
| FXVIXL2 | -0.012 * | -0.007 * | 0.031 ** | 0.081 *** | 0.112 *** | 0.140 *** | 0.167 *** | 38.29 | n.a. | 45.14 |
| AVEL1 | 0.025 *** | 0.011 *** | 0.020 *** | 0.031 *** | 0.037 *** | 0.042 *** | 0.040 *** | 55.43 | n.a. | 33.14 |
| AVEL2 | 0.025 *** | 0.012 *** | 0.017 *** | 0.032 *** | 0.041 *** | 0.046 *** | 0.041 *** | 41.14 | n.a. | 40.57 |

Notes: [a] Significance levels of MDM/MCW tests. One-sided tests in favor of indicated model: *** 1%, ** 5%, * 10%; one-sided tests in favor of random walk model: +++ 1%, ++ 5%, + 10%.
[b] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable.

second and fifth lags. For $h \leq 5$, it achieves better $\tilde{R}_h^2$ than the optimized ARIMA model and, in the aggregate, outperforms the random walk for $h \leq 2$. Adding the day-of-week dummies further improves the short-term forecasting performance of the AR*(5) model. The ARI$^S$($p$,$d$,$P_s$) model approaches this day-of-week phenomenon from a seasonal-AR perspective. Whether in optimized form or with a predetermined lags (listed as ARI$^S$(2,1,1) form, the short-term forecasting performance falls short of the AR*(5) model—but the long-term ($h \geq 5$) is virtually identical.

The third group of models in Table 2 are factor models, augmented in a few cases with day-of-week dummy variables. The CFC3 model reduces $\Delta y_{1t}, \dots, \Delta y_{25,t}$ to just three factors,[8] which capture 93.8%[9] of the variation in 25 variables. At $h = 1$, forecasts are 1 percent better than the those of the random walk model, but at longer horizons, this model is no longer competitive. The difference between this CFC3 model and random walk is not statistically significant, even though the model follows a very different strategy. Adding day-of-week dummies does make a material difference for short-horizon forecasts: the $\tilde{R}_h^2$ is statistically significantly better than the random walk and indicates that the one- and two-day sum of squared forecast errors are 6.2% and 3.6% lower. For medium and long-term forecasts, this advantage dissipates: CFC3D is equivalent to RW.

The CFL$k$ models capture the level values of the option volatility series. Table 2 lists results for three and ten factors. Though $k = 3$ is the best choice according to the elbow method (three factors capture 99.5% of the variation), the short-term forecasting performance is still relatively poor, and the long-term performance is only marginal better than for $k = 1$. Adding more factors captures more trends and improves the forecasting performance until, roughly, $k = 10$. These 10 factors capture 99.99% of the variation in the option volatility series. $\tilde{R}_h^2 > 0$ for all forecast horizons and statistically significantly so at $h = 1$; for 28% of the 175 forecasted series, this CFL10 is statistically preferred to the random walk.

The GG model is not competitive. As discussed before, the GG model estimates the shape of the volatility surface and predicts the evolution of this shape over the forecast horizon. It helps to run separate forecasts for the gap between the predicted and actual volatility surface at each separate maturity-delta combination—without it, the $\tilde{R}_h^2$ values would be $-1.079$ for $h = 1$ and $-0.045$ for $h = 25$—but even then, the random walk model is still statistically preferred for 15.4% of the 175 forecasted series.

Among the diffusion index models, the DI$k$FD version is most successful. The DI models extract factors from the five-day history of all option volatility series. Table 2 illustrates forecasting performance for $k = 5$ (strongly indicated by the elbow method and consistent with the notion that each lag may be represented by one factor, as in the CFC$k$ model discussed above) and $k = 10$. Increasing $k$ improves short-term performance ($h = 1$), but this gain evaporates when the forecast horizon lengthens. Adding other financial variables yields very little gain; adding the day-of-week dummy variables makes a substantial contribution in terms of short-term forecasting performance ($h \leq 2$).

The fourth group of models (bottom six rows of Table 2) focus more directly on the contribution of the day of week and the filter variable of FXVIX. Adding $\Delta f_{xt}$ (in the FXVIXC model) generates no improvement on AR*(5)D, similar to what was found with the diffusion index models. The same is true when $\Delta \bar{y}_t$ is added (i.e., the AVEC model). These variables add much to the regression models but bring no gain to forecasting.

The last four rows align more closely with the strong cross-sectional dependence concept: common factors are modeled respectively as $f_{xt}$ (the log of the FXVIX index that measures volatility in the overall currency markets) or as $\bar{y}_t$ (equivalent to the log of the geometric average of the 25 currency

[8] Three factors suffice on basis of the elbow method of selecting factors. The first factor captures 89.7% of the variation and would already be sufficient according to the eigenvalue rule of thumb. The CFC1 model performs similarly to the CFC3 model.

[9] Statistical descriptions of the model such as these pertain to the performance of the model over the entire 2010-2012 time period used to estimate parameters that are used for forecasting, i.e., all observations that are part of any estimation window.

options); the two models for each alternative differ in the way that this factor and the remaining filtered $\tilde{y}_{it}$ are forecast. In forecasting, both approaches yield significant improvements over the random walk model at all forecast horizons, especially in the long run ($h = 25$) where the mean squared prediction error is up to 19 percent lower.

CFL$k$, FXVIXL and AVEL models are all common factor models: why do they yield a different forecasting performance? As mentioned, $f_{xt}$ and $\bar{y}_t$ can be thought of as filters or "observable common factors" that are determined either externally (as with $f_{xt}$) or internally but deterministically (as with $\bar{y}_t$). In contrast, CFL$k$ models find common factors internally with a data-driven algorithm that generates factors both optimally and endogenously. As the CFL factors are designed to capture the maximal proportion of variation in the data, it may actually be surprising that the FXVIXL and AVEL models outperform the CFL$k$ models. However, first, note that the models are not nested. Forecasts based on CFL$k$ rely on forecasts of the factors and ignore any remaining patterns in the residuals. In contrast, the FXVIXL and AVEL models rely on forecasts of both the observable common factor and the filtered $\tilde{y}_{it}$. As the GG model shows, such residuals do carry useful information. Second, the five factors of the CFL5 model explain 96.2% of the variation in $f_{xt}$ and over 99.99% of the variation in $\bar{y}_t$. This shows that the latent common factors capture nearly all of the variation in the observable common factors—but the latter carry that information in a parsimonious fashion. Interestingly, there appears to be no gain in including both $f_{xt}$ and $\bar{y}_t$ as these variables carry very similar informational content. Third, the latent common factors are data-driven and thus endogenous. That is, a large shock throws off the correlation matrix, its eigenvectors, and thus the derived latent common factors. Such estimation noise is not present in observable common factors. It is true that $\bar{y}_t$ is also impacted by that large shock because some $y_{it}$ are subject to that shock, but the weights given to each $y_{it}$ in computing $\bar{y}_t$ remain equal to 1/25 for each $i$; those weights do not vary as they do for the latent common factors. The same may be said for $f_{xt}$ in relation to the broader foreign exchange options market.

### *5.2. An in-depth assessment*

Of course, Table 2 present a highly aggregative picture of the performance of these models. For a more disaggregate examination, consider four models that, according to Table 2, appear to be most successful: CFL10, DI10FD, AVEC, and FXVIXL1. Figures 3 and 4 show the percentage difference in the root mean squared prediction error ($RMSPE_{ih}$) of a given model relative to the random walk model, by forecast horizon, delta (measured horizontally) and maturity (measured vertically). This percentage difference is related to the out-of-sample $R^2_{ih}$:

$$\%D(RMSPE_{ih}) = 100\left(\left(1 - R^2_{ih}\right)^{1/2} - 1\right) = 100\left(\frac{RMSPE_{ih} - RMSPE^{rw}_{ih}}{RMSPE^{rw}_{ih}}\right)$$

It indicates more clearly the potential advantage of a model in generating forecasts. If a model generates smaller (greater) forecast errors, we have $R^2_{ih} > 0$ ($R^2_{ih} < 0$) and thus $\%D(RMSPE_{ih}) < 0$ ($> 0$). I.e., $R^2_{ih} = 0.05$ yields $\%D(RMSPE_{ih}) = -2.53$, implying that, for option $i$ at forecast horizon $h$, the comparison model generates an $RMSPE_{ih}$ that is 2.53% lower than the random walk model. Section 3.3 described tests of $R^2_{ih}$, which are therefore tests of $\%D(RMSPE_{ih})$ as well. In Figures 3 and 4, the statistical significance of a model that performs better than the random walk model is indicated by a color fading from dark brown (1% significance) to light yellow (20% significance) with significance levels of 10% or less highlighted in boldface font. In Figure 3, when the random walk model performs better, statistical significance is indicated by italicized font and a color fading from dark green (1%) to light green (20%) with significance levels of 10% or less highlighted in boldface font.

**Figure 3**: Percentage difference in RMSPE relative to random walk,
by forecast horizon h, delta and maturity (euro, 2010-2012)

A: DI10FD model

| | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -6.19 | -6.42 | -6.60 | -7.31 | -6.98 |
| | 2mo | -2.09 | -1.33 | -0.92 | -0.62 | -0.70 |
| | 3mo | -2.64 | -2.35 | -2.05 | -1.21 | -1.18 |
| | 6mo | -2.36 | -1.79 | -1.36 | -1.15 | -1.93 |
| | 12mo | -2.05 | -1.54 | -1.88 | -1.60 | -2.00 |
| h=2 | 1mo | -3.32 | -3.32 | -4.02 | -4.92 | -4.65 |
| | 2mo | 0.88 | 1.24 | 1.35 | 1.17 | 1.04 |
| | 3mo | -0.60 | -0.17 | 0.19 | 0.46 | 0.46 |
| | 6mo | 0.21 | 0.52 | 0.16 | 0.19 | -0.26 |
| | 12mo | 0.28 | 0.71 | -0.29 | -0.04 | -0.56 |
| h=5 | 1mo | 2.38 | 3.31 | 3.33 | 2.94 | 2.48 |
| | 2mo | 3.5 | 3.84 | 3.7 | 2.64 | 2.49 |
| | 3mo | 1.43 | 2.12 | 1.76 | 2.22 | 2.13 |
| | 6mo | 3.17 | 2.86 | 1.50 | 1.91 | 1.43 |
| | 12mo | 3.87 | 3.24 | 1.66 | 1.36 | 0.70 |
| h=10 | 1mo | 1.97 | 2.38 | 2.47 | 2.37 | 2.28 |
| | 2mo | 1.96 | 2.16 | 2.20 | 1.85 | 1.81 |
| | 3mo | 1.33 | 1.69 | 1.77 | 1.49 | 1.38 |
| | 6mo | 1.97 | 2.00 | 1.25 | 0.77 | 0.42 |
| | 12mo | 2.53 | 2.38 | 1.75 | 0.35 | 0.26 |
| h=15 | 1mo | 2.23 | 2.62 | 2.70 | 2.57 | 2.45 |
| | 2mo | 2.10 | 2.38 | 2.38 | 1.89 | 1.86 |
| | 3mo | 1.51 | 1.92 | 1.88 | 1.34 | 1.24 |
| | 6mo | 1.85 | 1.97 | 1.29 | 0.54 | 0.13 |
| | 12mo | 2.47 | 2.33 | 1.85 | 0.07 | -0.31 |
| h=20 | 1mo | 2.48 | 2.83 | 2.92 | 2.72 | 2.61 |
| | 2mo | 2.32 | 2.62 | 2.60 | 1.90 | 1.86 |
| | 3mo | 1.82 | 2.26 | 2.29 | 1.17 | 1.09 |
| | 6mo | 1.84 | 1.94 | 1.18 | 0.06 | -0.47 |
| | 12mo | 2.34 | 2.23 | 1.88 | -0.27 | -0.91 |
| h=25 | 1mo | 2.50 | 2.87 | 2.90 | 2.77 | 2.78 |
| | 2mo | 2.40 | 2.70 | 2.65 | 2.16 | 2.13 |
| | 3mo | 1.86 | 2.17 | 2.19 | 1.43 | 1.22 |
| | 6mo | 1.58 | 1.67 | 1.10 | -0.01 | -0.81 |
| | 12mo | 1.95 | 1.78 | 2.32 | -0.21 | -1.81 |

B: CFL10 model

| | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -1.48 | -2.05 | -2.54 | -1.74 | -1.97 |
| | 2mo | -0.98 | -0.24 | -1.73 | -0.50 | -0.91 |
| | 3mo | -1.99 | -1.81 | -3.01 | -0.07 | -1.51 |
| | 6mo | 0.80 | 1.24 | -1.75 | 0.12 | -3.75 |
| | 12mo | -3.29 | 0.21 | -1.64 | 0.83 | -1.26 |
| h=2 | 1mo | -0.60 | -1.53 | -2.62 | -1.99 | -2.15 |
| | 2mo | 1.55 | 1.53 | 0.19 | 1.45 | 0.82 |
| | 3mo | -0.58 | -0.55 | -1.83 | 0.63 | -0.81 |
| | 6mo | 1.88 | 1.88 | -1.16 | 1.08 | -2.53 |
| | 12mo | -2.54 | 0.81 | -1.28 | 1.96 | 0.32 |
| h=5 | 1mo | 0.15 | -0.73 | -1.95 | -1.77 | -2.53 |
| | 2mo | -0.02 | -0.28 | -1.85 | -1.48 | -2.67 |
| | 3mo | -0.91 | -0.20 | -1.57 | -0.22 | -1.96 |
| | 6mo | 1.05 | 1.48 | -0.11 | 2.08 | -0.92 |
| | 12mo | -2.29 | 0.09 | -0.63 | 3.27 | 1.96 |
| h=10 | 1mo | -0.06 | -1.31 | -2.91 | -3.68 | -4.48 |
| | 2mo | 0.41 | -0.24 | -2.11 | -2.03 | -3.28 |
| | 3mo | -0.62 | -0.29 | -2.02 | -1.15 | -2.41 |
| | 6mo | 1.26 | 1.81 | 0.61 | 3.11 | 1.04 |
| | 12mo | -1.96 | 0.33 | -0.18 | 4.31 | 5.00 |
| h=15 | 1mo | 0.80 | -0.59 | -2.50 | -3.71 | -4.63 |
| | 2mo | 0.52 | -0.42 | -2.44 | -2.68 | -3.76 |
| | 3mo | -0.13 | 0.11 | -1.57 | -1.00 | -2.09 |
| | 6mo | 1.19 | 1.87 | 1.03 | 3.48 | 2.34 |
| | 12mo | -1.85 | 0.08 | -0.06 | 4.39 | 6.55 |
| h=20 | 1mo | 1.32 | -0.21 | -2.16 | -3.74 | -4.78 |
| | 2mo | 0.39 | -0.72 | -2.63 | -3.24 | -4.05 |
| | 3mo | -0.17 | 0.01 | -1.55 | -1.52 | -2.11 |
| | 6mo | 1.15 | 1.82 | 1.20 | 3.39 | 3.30 |
| | 12mo | -2.03 | -0.45 | -0.61 | 3.89 | 7.55 |
| h=25 | 1mo | 1.80 | 0.11 | -1.76 | -3.54 | -4.64 |
| | 2mo | 0.13 | -1.19 | -2.89 | -3.57 | -4.18 |
| | 3mo | -0.05 | -0.09 | -1.46 | -1.50 | -2.06 |
| | 6mo | 1.54 | 2.26 | 1.66 | 3.92 | 4.34 |
| | 12mo | -1.60 | -0.35 | -0.74 | 3.56 | 7.78 |

Statistical significance is coded by color and font thickness and italics:

| | | | | |
|---|---|---|---|---|
| In favor of indicated model: | 1% | 5% | 10% | 20% |
| In favor of random walk model: | 1% | 5% | 10% | 20% |

Figure 3 examines DI10FD (a model of the first difference of log implied volatility) and CFL10 (a model of log implied volatility in levels). DI10FD does very well at very short-horizon forecasts ($h = 1$), providing gains of around 7% for 1-month options across the board and gains of around 1-to-2 percent for longer-maturity options. However, for $h = 5$, the random walk model has the clear upper hand with gains between 1 and 4 percent. This advantage persists for longer forecast horizons, even if only weakly statistically significant (or not at all).

In Table 2, CFL10 did not acquire many stars, although at each forecast horizon $\tilde{R}_h^2$ was positive. Figure 3, Panel B, shows why: (i) at each $h$, forecasts for some options are more accurate with CFL10 and for other options with random walk; (ii) at $h = 1$, CFL10 generally does better than random walk but gains are small; (iii) $h \geq 5$, forecasts for put options (delta less than 50) of random walk and CFL10 are equivalent, but there is an interesting divergence for put options: short-term call options (three months or less) are better handled by CFL10 forecasts whereas long-term call options are better forecast with the random walk model. For example, for $h = 20$ and delta $= 90$, this divergence runs as high as a 4.78% gain for CFL10 forecasts of 1-month call options and a 7.55% gain for random walk forecasts of 12-month options.

Figure 4 turns to models with observable common factors. AVEC models the first-difference of log implied volatility; FXVIXL1 aims for log implied volatility in levels. Note that random walk model is nested within AVEC and FXVIXL1. Thus, tests can only reject random walk in favor of the alternative model; they cannot reject the alternative model in favor of random walk. Moreover, $\%D(RMSPE_{ih})$ can be positive, which seems to favor random walk, and yet tests can indicate rejection of the random walk model.

Again, the first-difference AVEC model generates strong forecasting performance at short horizons $h \leq 2$, with $\%D(RMSPE_{ih})$ favoring the random walk model for $h \geq 5$. On the other hand the FXVIXL1 model does well for virtually all options for longer forecast horizons with gains running as high as 18.67%, while still testing favorably for short horizons ($h \leq 2$) as well. One weak spot is the 2-month call option for delta $= 90$ at all horizons. This performance is incongruous with both 2-month call option for delta $= 75$ and 1- and 3-month call options for delta $= 90$.[10]

***5.3. Robustness: another time period, another currency***

An important question is whether the patterns shown in Section 4 extend to other time periods and other currencies. This section examines euro options in 2013-2015 as well as Canadian dollar options in both 2010-2012 and 2013-2015. In this, the main models to be highlighted are ARIMA, CFL10, DI10FD, the AR2p5D day-of-week model, and the models involving $f_{xt}$ and $\bar{y}_t$. A few other models are added if they performed particularly well. The aggregate results of all models are provided in Appendix A.1, together with a few of the detailed results that parallel those in Section 5.2.

[10] Figure A.1 in Appendix A.1.1 shows the disaggregate performance of ARIMA. Table 2 indicated that ARIMA forecasts dominates random walk forecasts in the aggregate for $h = 1$ but are equivalent and possibly slightly worse for $h \geq 2$. Figure A.1 presents a more nuanced picture. For example, ARIMA outperforms random walk for about half of the options for $h \geq 15$ though with small gains, but shortfalls for the other forecasts in this group outweigh the gains. ARIMA also performed better for put (delta < 50) options than call (delta > 50) options.

**Figure 4**: Percentage difference in RMSPE relative to random walk,
by forecast horizon h, delta and maturity (euro, 2010-2012)

A: AVEC model

| h | maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -6.57 | -7.26 | -7.78 | -8.46 | -7.48 |
| | 2mo | -1.13 | -0.70 | -0.83 | -0.76 | -0.37 |
| | 3mo | -1.98 | -1.45 | -1.66 | -1.51 | -1.51 |
| | 6mo | -1.72 | -1.27 | -1.50 | -1.64 | -1.88 |
| | 12mo | -1.75 | -1.39 | -2.14 | -1.80 | -1.91 |
| h=2 | 1mo | -4.35 | -5.09 | -5.87 | -6.69 | -6.04 |
| | 2mo | -0.01 | 0.14 | 0.05 | -0.15 | -0.03 |
| | 3mo | -0.56 | -0.37 | -0.66 | -0.94 | -0.78 |
| | 6mo | -0.54 | -0.26 | -0.45 | -0.71 | -0.63 |
| | 12mo | -0.44 | -0.5 | -0.93 | -1.14 | -1.01 |
| h=5 | 1mo | 0.01 | -0.02 | 0.03 | 0.24 | 0.43 |
| | 2mo | 0.23 | 0.3 | 0.22 | 0.21 | 0.28 |
| | 3mo | 0.34 | 0.36 | 0.14 | -0.09 | -0.18 |
| | 6mo | 0.3 | 0.38 | 0.2 | -0.14 | -0.26 |
| | 12mo | 0.59 | 0.54 | 0.35 | -0.05 | -0.41 |
| h=10 | 1mo | 0.75 | 0.75 | 0.84 | 1.06 | 1.28 |
| | 2mo | 0.84 | 0.8 | 0.7 | 0.71 | 0.79 |
| | 3mo | 0.87 | 0.76 | 0.49 | 0.25 | 0.19 |
| | 6mo | 0.79 | 0.7 | 0.32 | -0.19 | -0.47 |
| | 12mo | 1.15 | 0.94 | 0.47 | -0.31 | -0.91 |
| h=15 | 1mo | 1.14 | 1.19 | 1.3 | 1.48 | 1.64 |
| | 2mo | 1.22 | 1.21 | 1.16 | 1.16 | 1.25 |
| | 3mo | 1.24 | 1.19 | 0.98 | 0.71 | 0.62 |
| | 6mo | 1.07 | 0.94 | 0.51 | -0.16 | -0.59 |
| | 12mo | 1.54 | 1.24 | 0.66 | -0.37 | -1.31 |
| h=20 | 1mo | 1.45 | 1.58 | 1.72 | 1.91 | 2.05 |
| | 2mo | 1.51 | 1.53 | 1.51 | 1.49 | 1.56 |
| | 3mo | 1.49 | 1.45 | 1.28 | 0.97 | 0.81 |
| | 6mo | 1.28 | 1.1 | 0.6 | -0.28 | -0.9 |
| | 12mo | 1.78 | 1.47 | 0.81 | -0.52 | -1.8 |
| h=25 | 1mo | 1.88 | 2.09 | 2.3 | 2.47 | 2.59 |
| | 2mo | 1.83 | 1.93 | 1.99 | 2.01 | 2.08 |
| | 3mo | 1.67 | 1.71 | 1.64 | 1.36 | 1.19 |
| | 6mo | 1.29 | 1.18 | 0.76 | -0.12 | -0.83 |
| | 12mo | 1.76 | 1.47 | 0.86 | -0.56 | -2.06 |

B: FXVIXL1 model

| h | maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -2.18 | -2.34 | -2.55 | -2.65 | -2.26 |
| | 2mo | -0.58 | -0.66 | -1.11 | -0.86 | 0.79 |
| | 3mo | -0.55 | -0.44 | -1.00 | -0.40 | -0.12 |
| | 6mo | 0.05 | 0.61 | 0.83 | 1.07 | 0.98 |
| | 12mo | -0.16 | 0.62 | 1.16 | 2.39 | 1.42 |
| h=2 | 1mo | -2.10 | -2.64 | -3.47 | -4.02 | -3.76 |
| | 2mo | -0.72 | -0.56 | -0.93 | -0.71 | 1.68 |
| | 3mo | -1.85 | -1.53 | -1.70 | -1.21 | -0.87 |
| | 6mo | -2.01 | -1.41 | -1.16 | -0.50 | -0.32 |
| | 12mo | -1.93 | -1.63 | -0.94 | -0.05 | -0.42 |
| h=5 | 1mo | -2.01 | -2.64 | -3.92 | -5.05 | -6.03 |
| | 2mo | -2.88 | -3.14 | -4.54 | -4.82 | 0.14 |
| | 3mo | -2.63 | -2.71 | -3.90 | -3.41 | -2.78 |
| | 6mo | -3.10 | -2.78 | -3.02 | -1.94 | -1.26 |
| | 12mo | -3.40 | -3.13 | -2.04 | -0.50 | -0.24 |
| h=10 | 1mo | -3.85 | -5.38 | -8.12 | -10.72 | -12.91 |
| | 2mo | -4.47 | -5.43 | -8.36 | -9.89 | -1.52 |
| | 3mo | -4.17 | -4.38 | -6.82 | -7.05 | -6.15 |
| | 6mo | -4.50 | -4.36 | -5.36 | -3.89 | -2.88 |
| | 12mo | -4.92 | -4.66 | -3.98 | -1.80 | -1.25 |
| h=15 | 1mo | -4.78 | -6.71 | -9.99 | -13.10 | -15.74 |
| | 2mo | -5.44 | -6.69 | -10.13 | -12.06 | -2.43 |
| | 3mo | -5.08 | -5.29 | -8.00 | -8.37 | -7.54 |
| | 6mo | -5.57 | -5.39 | -6.55 | -4.69 | -3.58 |
| | 12mo | -6.11 | -6.13 | -5.40 | -3.09 | -2.44 |
| h=20 | 1mo | -6.06 | -8.39 | -11.77 | -15.10 | -17.82 |
| | 2mo | -6.80 | -8.16 | -11.76 | -13.85 | -2.76 |
| | 3mo | -6.09 | -6.39 | -9.21 | -9.60 | -8.36 |
| | 6mo | -6.67 | -6.45 | -7.49 | -5.60 | -4.09 |
| | 12mo | -7.40 | -8.04 | -7.09 | -4.40 | -2.42 |
| h=25 | 1mo | -8.06 | -10.44 | -13.44 | -16.39 | -18.67 |
| | 2mo | -8.77 | -10.25 | -13.68 | -15.30 | -3.65 |
| | 3mo | -7.47 | -7.99 | -10.86 | -10.97 | -9.21 |
| | 6mo | -7.73 | -7.63 | -8.66 | -6.35 | -4.43 |
| | 12mo | -8.52 | -9.46 | -8.59 | -5.31 | -2.79 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| 1% | 5% | 10% | 20% |
|---|---|---|---|

Table 3 starts off with the euro in 2013-2015. In brief summary: (i) ARIMA does provide some improvement over the random walk, especially in 1-month and 3-month options, but the gain in the RMSPE is always less than 1 percent. (ii) In the aggregate, CFL10 does not distinguish itself from the random walk, but this hides the fact that the performance is highly variable: CLF10 is preferred for 22.3% of the 175 option/horizon combinations and random walk forecasts perform better for 20.57% of them. As in 2010-2012, there is no clear pattern in this by forecast horizon. (iii) DI10FD fails to provide better short-horizon forecasts as it did before, and it yields to the random walk model for longer horizons. (iv) AVEC has the same performance as before, relying as it does on day-of-week effects. (v) FXVIXL models are not nearly as effective as in 2013-2015. In fact, AVEL models do better.

**Table 3**: Aggregate out-of-sample R2 and MDM/MCW tests, and summary of separate MDM/MCW and sign tests: Euro 2013-2015

| Type of | Forecast horizon | | | | | | | Share preferred[a] | | Sign |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Model | R.Walk | test |
| ARIMA | -0.003 * | -0.006 * | -0.003 | 0.002 * | 0.006 ** | 0.005 *** | 0.006 *** | 22.86 | n.a. | 5.71 |
| CFL10 | 0.007 | 0.009 | 0.008 | 0.012 | 0.014 | 0.009 | 0.011 | 22.29 | 20.57 | 22.29 |
| DI10FD | 0.022 | -0.003 | -0.026 | -0.018 | -0.023 | -0.029 | -0.043 + | 1.71 | 29.14 | 5.71 |
| AR*(5) | 0.005 ** | 0.005 * | -0.001 | -0.009 | -0.017 | -0.027 | -0.036 | 10.29 | n.a. | 1.71 |
| AR*(5)D | 0.027 *** | 0.016 *** | 0.002 | -0.001 | -0.012 | -0.019 | -0.029 | 16.57 | n.a. | 7.43 |
| FXVIXC | 0.016 *** | 0.001 ** | -0.006 | -0.014 | -0.022 | -0.033 | -0.042 | 12.57 | n.a. | 6.29 |
| AVEC | 0.028 *** | 0.015 *** | -0.002 | -0.003 | -0.012 | -0.019 | -0.029 | 14.86 | n.a. | 9.71 |
| FXVIXL1 | -0.046 | -0.068 | -0.069 | -0.080 | -0.053 | -0.021 | 0.014 | 28.57 | n.a. | 7.43 |
| FXVIXL2 | -0.042 * | -0.037 | -0.042 | -0.033 | -0.015 * | 0.002 * | 0.016 * | 29.71 | n.a. | 4.57 |
| AVEL1 | 0.019 | 0.018 | 0.008 | 0.011 | 0.015 | 0.016 | 0.021 | 41.71 | n.a. | 16.00 |
| AVEL2 | 0.015 | 0.014 | 0.005 | 0.005 | 0.008 | 0.008 | 0.014 | 40.00 | n.a. | 17.71 |

Notes: [a] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable. Significance levels: one-sided tests in favor of indicated model: *** 1%, ** 5%, * 10%; one-sided tests in favor of random walk model: +++ 1%, ++ 5%, + 10%.

Table 4 gives an overview of the forecast performance in the case of Canadian dollar options, for both 2010-2012 and 2013-2015. ARIMA does not generate much improvement in the aggregate. However, as shown in Appendix A.1.2, for 2010-2012, ARIMA predicts call (delta $\geq$ 50) options better than random walk at all $h$ but does substantially worse on put (delta < 50) options. For 2013-2015, ARIMA is doing well only for $h=1$. CFL10 offers a mixed bag once again, in both periods. In 2013-2015, CFC3D generates a strong performance for $h \leq 2$, in line with the AR2p5D models even though it follows a different strategy. I.e., day-of-week effects are substantial, but they do not help for long-horizon forecasts. The DI10FD model shows some ability to forecast over short horizons ($h \leq 2$) only. Similarly, for $h \leq 2$, AVEC and FXVIXC also provide more accurate forecasts than the random walk model. FXVIXL1 does well for many (60.8%) option/horizon combinations in 2010-2012 but flourishes only for short-horizon forecasts in 2013-2015. The FXVIXL1 model perform better than AVEL1 in 2010-2012 but worse in 2013-2015; this is not only in the aggregate but also for just about each Canadian dollar option (Figures A.3 and A.4 in Appendix A.1.2).

**Table 4**: Aggregate out-of-sample R2 and MDM/MCW tests, and summary of separate MDM/MCW and sign tests: Canadian Dollar

| Type of Model | Forecast horizon $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Share preferred[a] Model | R.Walk | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| A: 2010-2012 | | | | | | | | | | |
| ARIMA | -0.006 | -0.010 | -0.003 | -0.004 | -0.004 | -0.005 | -0.005 | 33.71 | n.a. | 22.86 |
| CFL10 | 0.011 | 0.005 | 0.000 | -0.008 | -0.009 | -0.008 | -0.007 | 4.57 | 1.71 | 5.14 |
| DI10FD | 0.022 | -0.018 | -0.007 | 0.016 | 0.009 | 0.020 | 0.033 | 4.57 | 2.29 | 57.71 |
| AR*(5) | -0.006 * | -0.004 | 0.011 | 0.015 | 0.014 | 0.023 | 0.029 | 17.14 | n.a. | 69.14 |
| AR*(5)D | 0.024 *** | 0.033 *** | 0.000 | 0.007 | 0.009 | 0.018 | 0.028 | 26.29 | n.a. | 70.29 |
| FXVIXC | 0.058 *** | 0.061 *** | 0.025 ** | 0.025 | 0.024 | 0.028 | 0.035 | 52.57 | n.a. | 68.57 |
| AVEC | 0.034 *** | 0.037 *** | -0.004 | 0.005 | 0.008 | 0.017 | 0.027 | 30.29 | n.a. | 69.71 |
| FXVIXL1 | 0.019 | 0.035 | 0.023 | 0.033 | 0.046 | 0.071 | 0.095 | 60.57 | n.a. | 61.14 |
| FXVIXL2 | -0.008 ** | -0.028 * | -0.037 | -0.021 | 0.003 | 0.030 | 0.054 | 15.43 | n.a. | 5.71 |
| AVEL1 | 0.014 | 0.010 | 0.015 | 0.015 | 0.015 | 0.014 | 0.014 | 48.57 | n.a. | 31.43 |
| AVEL2 | 0.012 | 0.009 | 0.011 | 0.013 | 0.014 | 0.013 | 0.011 | 43.43 | n.a. | 27.43 |
| B: 2013-2015 | | | | | | | | | | |
| ARIMA | 0.039 *** | 0.007 *** | -0.009 | -0.009 | -0.004 | -0.003 | -0.006 | 30.29 | n.a. | 11.43 |
| CFC3D | 0.159 *** | 0.070 *** | -0.046 +++ | -0.067 ++ | -0.061 | -0.065 | -0.089 | 19.43 | 22.86 | 17.71 |
| CFL10 | 0.096 *** | 0.047 *** | -0.010 | -0.014 | 0.002 | 0.001 | -0.008 | 17.71 | 12.00 | 21.71 |
| DI10FD | 0.150 *** | 0.054 * | -0.014 | -0.012 | -0.015 | -0.046 | -0.069 | 10.86 | 1.71 | 12.57 |
| AR*(5) | 0.055 *** | 0.006 ** | -0.046 | -0.065 | -0.060 | -0.064 | -0.079 | 24.57 | n.a. | 9.71 |
| AR*(5)D | 0.150 *** | 0.077 *** | -0.023 | -0.040 | -0.042 | -0.051 | -0.070 | 27.43 | n.a. | 10.86 |
| FXVIXC | 0.129 *** | 0.060 *** | -0.025 | -0.048 | -0.047 | -0.054 | -0.073 | 22.86 | n.a. | 8.57 |
| AVEC | 0.162 *** | 0.082 *** | -0.021 | -0.037 | -0.039 | -0.049 | -0.069 | 28.57 | n.a. | 12.00 |
| FXVIXL1 | 0.015 | -0.021 | -0.035 | -0.090 | -0.144 | -0.148 | -0.135 | 8.00 | n.a. | 1.14 |
| FXVIXL2 | 0.005 ** | 0.006 ** | -0.015 | -0.048 | -0.094 | -0.099 | -0.095 | 21.71 | n.a. | 6.86 |
| AVEL1 | 0.095 | 0.046 | -0.010 | -0.009 | 0.006 | 0.006 | -0.001 | 34.29 | n.a. | 16.57 |
| AVEL2 | 0.089 | 0.037 | -0.013 | -0.013 | 0.001 | 0.003 | -0.004 | 34.29 | n.a. | 12.57 |

Notes: [a] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable. Significance levels: one-sided tests in favor of indicated model: *** 1%, ** 5%, * 10%; one-sided tests in favor of random walk model: +++ 1%, ++ 5%, + 10%.

## 6. Trading the IV Forecasts

### *6.1 Trading strategy*

To evaluate the economic profit of trading the IV forecasts, this paper engages in a trading strategy similar to Chalamandaris and Tsekrekos (2011, 2014), which seeks to identify a portfolio of delta-hedged call option contracts with the strongest buy or sell signal based on the implied volatility forecasts. A brief description will be given below.

For a given forecast horizon $h$, $\hat{d}_{ht}(i) = R[\hat{\sigma}_{i,t+h} - \sigma_{i,t}]$ stands for the ranking of the difference between the $h$-day-ahead IV forecast and the current IV for option $i$, in an ascending order.[11] Let $P_{ht}^{-}$ and $P_{ht}^{+}$ be the set of options with the $p$ lowest and $p$ highest forecasted deviations:

$$P_{ht}^{-} = arg_i\{\hat{d}_{ht}(i) \leq p\} \tag{6.1}$$

$$P_{ht}^{+} = arg_i\{\hat{d}_{ht}(i) \geq N - p\} \tag{6.2}$$

Let $A_{ht}^{-}$ and $A_{ht}^{+}$ be the average deviation among the options in $P_{ht}^{-}$ and $P_{ht}^{+}$:

$$A_{ht}^{-} = \frac{1}{p}\sum_{i \in P_{ht}^{-}}\left(\hat{\sigma}_{i,t+h} - \sigma_{i,t}\right) \tag{6.3}$$

$$A_{ht}^{+} = \frac{1}{p}\sum_{i \in P_{ht}^{+}}\left(\hat{\sigma}_{i,t+h} - \sigma_{i,t}\right) \tag{6.4}$$

For each day, a long-only or short-only trading strategy is implemented based on the values of $A_{ht}^{-}$ and $A_{ht}^{+}$. Note that $A_{ht}^{-} < A_{ht}^{+}$ by design. If $0 \leq A_{ht}^{-} < A_{ht}^{+}$, the trader would go long for the call option contracts in $P_{ht}^{+}$ because there is no sell signal from the IV forecasts. If $A_{ht}^{-} < A_{ht}^{+} \leq 0$, the trader would go short for the call option contracts in $P_{ht}^{-}$ because there is no buy signal from the forecasts. However, if $A_{ht}^{-} < 0 < A_{ht}^{+}$, the trader would make his/her decision based on the following rule: if $|A_{ht}^{+}| - |A_{ht}^{-}| > 0$, long $P_{ht}^{+}$; otherwise, short $P_{ht}^{-}$. Defining $I(L)$ be an indicator function, equal to 1 (0) if the logical argument $L$ is true (false), let $s_{ht}$ indicate the buy/sell position:

$$s_{ht} = 2I(|A_{ht}^{+}| > |A_{ht}^{-}|) - 1 \tag{6.5}$$

Thus, $s_{ht} = 1$ for a buy position and $= -1$ for a sell position. After the contracts and the position are determined, invest total amount of \$1000 cash into the $p$ selected contracts with the following portfolio weights:

$$w_{hi} = \frac{|\hat{\sigma}_{i,t+h} - \sigma_{i,t}|}{\sum_{j \in P_{ht}}|\hat{\sigma}_{j,t+h} - \sigma_{j,t}|}\ ,\ i \in P_{ht} \tag{6.6}$$

where $P_{ht}$ is either $P_{ht}^{+}$ or $P_{ht}^{-}$. Each long (short) call option position is delta-hedged by selling (buying) the underlying FX exchange rate. Each option position is held for one trading day and is closed on the next day. This is repeated for $h$ days, and the total $h$-day profit is accumulated. The daily profit earned through option $i$ is calculated as

$$\pi_{hi,t+1} = s_{ht}\left(\left(C_{i,t+1} - C_{i,t}\right) - \left(S_{t+1} - S_t\right)D_i - \frac{r_f}{360}a_t\left(C_{i,t} - D_i S_t\right)\right) \tag{6.7}$$

where $C_{i,t} = C(S_t, M_i, K_{it}, \sigma_{it} + s_{ht}b)$ is today's Black-Scholes call option price for opening the position, $C_{i,t+1} = C(S_{t+1}, M_i - 1, K_{it}, \sigma_{i,t+1} - s_{ht}b)$ is tomorrow's call option price for closing the position, $S_t$ is the FX spot rate, $M_i$ is time to maturity in days, $D_i$ is the delta call,[12] $r_f$ is the risk-free annual interest rate (set equal to 0.01, which is common in the literature but somewhat high relative to the estimates in Binsbergen et al (2019, Table 1)), $a_t$ is the number of calendar days between the two trading dates, $K_{it}$ is the strike price calculated on basis of the option delta, and $b$ is the transaction cost in basis points charged for trading options. The daily profit is turned into a daily rate of return by dividing it with $D_i S_t - C_{it} > 0$,

[11] Note that in earlier sections of this paper, $y_{it}$ and $\hat{y}_{it}$ denoted the realized and forecasted log implied volatility. Thus $y_{it} = \ln \sigma_{it}$ and $\hat{\sigma}_{it} = \exp(\hat{y}_{it})$.

[12] In this context, note that, e.g., our delta = 90 corresponds with a 10-delta call.

which is the value of the funds involved in the trade of the call option (Cao and Han, 2013; Cao et al., 2021):

$$r^c_{hi,t+1} = \frac{\pi_{hi,t+1}}{D_i S_t - C_{it}} \tag{6.8}$$

The daily return of the portfolio equals $r^c_{h,t+1} = \sum_i w_{hit} r^c_{hi,t+1}$, and the average daily return over the length of the trading period is denoted as $\bar{r}^c_{ht} = \sum_{j=1}^{h} r^c_{h,t+j}/h$. For reporting purposes, this is annualized by multiplying it with the number of trading days per year (252).

### 6.2. Assessing the trading strategies

The trading strategy is evaluated in two aspects: whether the average daily return is positive and whether the strategy is more profitable than a random walk strategy. For this, $\bar{r}^c_{ht}$ is computed for 252 days, separately for each $h$ and for selected forecasting models. If $h > 1$, $\bar{r}^c_{ht}$ is likely subject to serial correlation as $\bar{r}^c_{ht}, \bar{r}^c_{h,t+1}, \ldots, \bar{r}^c_{h,t+h}$ have trading days in common. Therefore, statistically, the test whether the expected value of $\bar{r}^c_{ht}$, denoted as $\mu_{\bar{r}^c_h}$, equals 0 is similar to the MDM/MCW test discussed in Section 4.2.

Trading on random walk forecasts cannot be done with this trading strategy, since options cannot be ranked by their IV forecast deviation as the random walk forecast $\hat{\sigma}_{i,t+h}$ is equal to $\sigma_{it}$. Instead, any portfolio of randomly selected option contracts with a randomly selected buy or sell position is consistent with the principle of the random walk. The expected value of the random walk strategy is the average over all possible portfolios and long/short positions. Since each option contract has an equal chance of being included in the portfolio, this amounts to populating the average portfolio with all 25 options with weights of $0.04$, computing $\bar{r}^{rw}_{ht}(s_{ht} = 1)$ for the long position and $\bar{r}^{rw}_{ht}(s_{ht} = -1)$ for the short position, and taking the simple average of these two values. Denote this average return as $\bar{r}^{rw}_{ht}$, and denote its expectation $E[\bar{r}^{rw}_{ht}]$ as $\mu_{\bar{r}^{rw}_h}$. The hypothesis that $\mu_{\bar{r}^c_h}$ equals $\mu_{\bar{r}^{rw}_h}$ is tested by examining whether the mean of $\bar{r}^c_{ht} - \bar{r}^{rw}_{ht}$ exceeds 0. Again, because of serial correlation, the proper statistical tool is the MDM/MCW test.

### *6.3. Results*

Section 5 concluded that, generally, the CFL10, DiFD10, and FXVIX- and AVE-related models (in first difference or in levels) yielded forecasts that are commendable in different ways. Thus, these models are examined as bases for the two trading strategies. Table 6 reports on only three models for different levels of trading costs ($b = 0$, $2$, and $5$ basis points), forecast horizons ($h = 1, 2, 5, \ldots, 25$ days), time periods (2010-2012 and 2013-2015), and both currencies (euro and Canadian dollar). The three models are selected as follows: one of CFL10 and DiDF10, one of FXVIXC and AVEC, and one of FXVIXL1, FXVIXL2, AVEL1 and AVEL2; within each subgroup, the most successful model is shown. In Appendix A.3, a full comparison is presented.

Table 6 reports average annualized returns. Significance levels of the test whether the return is positive are shown with superscripted "*" symbols; significance levels of the test whether profit is greater than the expected profit from a random walk strategy are shown with subscripted "+" symbols. For $b = 0$, the "+" symbols are omitted as the two tests overlap.

Without trading costs, many forecasting models beat the random walk strategy especially when relying on short-horizon ($h = 1$ or 2) forecasts. Statistically significant economic returns of 5 or 6 percent are feasible for the euro in both time periods. For AVEL2 in 2010-2012, returns to trading on longer forecast windows remain positive; for CFL10 and AVEC, they disappear. It is nearly the opposite in 2013-

**Table 6**: Returns to trading implied volatility

| | $b = 0$ | | | $b = 2$ | | | $b = 5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| A1: Euro, 2010-2012 | | | | | | | | | |
| $h$ | CFL10 | AVEC | AVEL2 | CFL10 | AVEC | AVEL2 | CFL10 | AVEC | AVEL2 |
| 1 | $5.48^{***}$ | $6.07^{***}$ | $5.21^{***}$ | $1.96^{*}_{+++}$ | $3.64^{**}_{+++}$ | $2.29^{*}_{+++}$ | $-3.30_{+++}$ | $0.00_{+++}$ | $-2.10_{+++}$ |
| 2 | $5.17^{***}$ | $4.64^{***}$ | $4.82^{***}$ | $1.52_{+++}$ | $1.95^{*}_{+++}$ | $1.58_{+++}$ | $-3.94_{+++}$ | $-2.09_{+++}$ | $-3.28_{+++}$ |
| 5 | 0.94 | $1.46^{*}$ | $3.26^{***}$ | $-2.77_{+++}$ | $-1.33_{+++}$ | $-0.24_{+++}$ | $-8.34_{+++}$ | $-5.51_{+++}$ | $-5.49_{+++}$ |
| 10 | -0.18 | 0.09 | $3.19^{***}$ | $-3.99_{+++}$ | $-3.14_{++}$ | $-0.52_{+++}$ | $-9.71_{+++}$ | $-7.98_{+++}$ | $-6.10_{+++}$ |
| 15 | -0.54 | 0.05 | $3.39^{***}$ | $-4.41_{++}$ | $-3.34_{+}$ | $-0.39_{+++}$ | $-10.21_{+++}$ | $-8.42_{+++}$ | $-6.07_{+++}$ |
| 20 | -0.74 | 0.18 | $3.40^{***}$ | -4.59 | $-3.28_{++}$ | $-0.52_{+++}$ | $-10.37_{+++}$ | $-8.46_{+++}$ | $-6.39_{+++}$ |
| 25 | -0.66 | 0.14 | $3.31^{***}$ | $-4.43_{+}$ | $-3.33_{++}$ | $-0.79_{+++}$ | $-10.08_{+++}$ | $-8.53_{+++}$ | $-6.93_{+++}$ |
| A2: Euro, 2013-2015 | | | | | | | | | |
| $h$ | DI10FD | FXVIXC | AVEL2 | DI10FD | FXVIXC | AVEL2 | DI10FD | FXVIXC | AVEL2 |
| 1 | $6.05^{**}$ | $5.57^{**}$ | $5.31^{**}$ | $3.25_{+++}$ | $2.37_{+++}$ | $2.46_{+++}$ | $-0.97_{+++}$ | $-2.44_{+++}$ | $-1.80_{+++}$ |
| 2 | 2.94 | $4.36^{**}$ | 1.82 | $0.12_{+++}$ | $1.52_{+++}$ | $-0.93_{+}$ | $-4.11_{+++}$ | $-2.73_{+++}$ | $-5.07_{+++}$ |
| 5 | -0.47 | 2.02 | 0.94 | -3.35 | $-1.11_{+}$ | $-1.75_{++}$ | $-7.66_{+++}$ | $-5.81_{+++}$ | $-5.79_{+++}$ |
| 10 | 1.50 | 4.59 | -0.05 | $-1.25_{+++}$ | $1.74_{+++}$ | -2.69 | $-5.37_{+++}$ | $-2.55_{+++}$ | $-6.64_{+++}$ |
| 15 | $2.91^{*}$ | $5.21^{**}$ | -0.78 | $0.22_{+++}$ | $2.49_{+++}$ | -3.41 | $-3.80_{+++}$ | $-1.59_{+++}$ | $-7.34_{+++}$ |
| 20 | $4.10^{**}$ | $5.94^{**}$ | -0.94 | $1.46_{+++}$ | $3.38^{*}_{+++}$ | -3.47 | $-2.49_{+++}$ | $-0.46_{+++}$ | $-7.26_{+++}$ |
| 25 | $4.87^{***}$ | $6.48^{***}$ | -0.55 | $2.25^{**}_{+++}$ | $3.99^{**}_{+++}$ | -3.12 | $-1.68_{+++}$ | $0.25_{+++}$ | $-6.97_{+++}$ |
| B1: Canadian dollar, 2010-2012 | | | | | | | | | |
| $h$ | CFL10 | FXVIXC | AVEL2 | CFL10 | FXVIXC | AVEL2 | CFL10 | FXVIXC | AVEL2 |
| 1 | $14.48^{***}$ | $11.18^{***}$ | $11.98^{***}$ | $7.40^{***}_{+++}$ | $6.04^{**}_{+++}$ | $5.28^{**}_{+++}$ | $-3.20_{+++}$ | $-1.66_{+++}$ | $-4.78_{+++}$ |
| 2 | $10.05^{***}$ | $8.83^{***}$ | $9.15^{***}$ | $2.87_{+++}$ | $3.47^{*}_{+++}$ | $2.66_{+++}$ | $-7.91_{++}$ | $-4.57_{+++}$ | $-7.07_{+++}$ |
| 5 | $6.99^{***}$ | $6.01^{**}$ | $4.36^{*}$ | $-1.33_{+}$ | $-1.92_{+}$ | $-1.80_{+}$ | -13.82 | -13.81 | -11.03 |
| 10 | $5.30^{**}$ | $3.37^{*}$ | $5.45^{*}$ | -3.15 | -3.99 | $-0.20_{+}$ | -15.83 | -15.04 | $-8.66_{+}$ |
| 15 | $3.44^{**}$ | 2.16 | $6.16^{*}$ | -4.95 | -4.87 | $0.55_{++}$ | -17.54 | -15.42 | $-7.87_{++}$ |
| 20 | 2.28 | 1.88 | 6.06 | -6.25 | -4.43 | $0.42_{++}$ | -19.05 | -13.91 | $-8.04_{++}$ |
| 25 | 1.49 | 1.65 | 5.91 | -6.97 | -4.22 | $0.22_{++}$ | -19.65 | -13.03 | $-8.32_{+}$ |
| B2: Canadian dollar, 2013-2015 | | | | | | | | | |
| $h$ | CFL10 | AVEC | AVEL2 | CFL10 | AVEC | AVEL2 | CFL10 | AVEC | AVEL2 |
| 1 | $3.48^{*}$ | 2.38 | $3.38^{*}$ | -2.27 | -3.35 | -2.38 | -10.92 | -11.95 | -11.02 |
| 2 | 2.82 | $3.67^{*}$ | 3.17 | -2.84 | -2.05 | -2.52 | -11.33 | -10.63 | -11.07 |
| 5 | 2.48 | 3.47 | $3.63^{*}$ | -3.06 | -2.26 | $-2.05_{+}$ | -11.37 | -10.85 | -10.57 |
| 10 | 2.13 | 2.38 | 2.76 | -3.34 | -3.25 | -2.90 | -11.55 | -11.70 | -11.39 |
| 15 | 1.72 | 2.02 | 2.15 | -3.65 | -3.60 | -3.49 | -11.72 | -12.04 | -11.94 |
| 20 | 1.90 | 2.20 | 2.25 | -3.45 | -3.41 | -3.43 | -11.49 | -11.84 | -11.96 |
| 25 | 1.94 | 2.13 | 2.18 | -3.37 | -3.49 | -3.49 | -11.36 | -11.93 | -12.01 |

Notes: Across all $h$, the average random-walk rate of return to trading in euro options equals approximately 0 ($b = 0$), $-5.45$ ($b = 2$), and $-13.63$ ($b = 10$) in 2010-2012, and 0 ($b = 0$), $-5.42$ ($b = 2$), and $-13.54$ ($b = 10$) in 2013-2015. The average random-walk rate of return to trading in CAD options equals approximately 0 ($b = 0$), $-5.41$ ($b = 2$), and $-13.53$ ($b = 10$) in 2010-2012, and 0 ($b = 0$), $-5.43$ ($b = 2$), and $-13.57$ ($b = 10$) in 2013-2015.
In columns with $b = 0$, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0 = \mu_{\bar{r}_h^{rw}}$ (*** 1%, ** 5%, * 10%). In columns with $b = 2$ and $b = 5$, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0$ (*** 1%, ** 5%, * 10%), whereas subscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = \mu_{\bar{r}_h^{rw}}$ (+++ 1%, ++ 5%, + 10%).

2015, where returns to trading on AVEL2 forecasts disappear for rising $h$ but returns to trading on DI10FD and FXVIXC first fade and then rise again as $h$ increases. For the Canadian dollar in 2010-2012, returns are as large as 14.48 percent and, as with the euro, fade with longer forecast windows. In 2013-2015, returns are much lower at roughly 3.5 percent for $h = 1$ and diminishing slightly as $h$ rises—and not particularly significantly different from 0.

With trading costs, these returns are necessarily lower. As an indication of the effect of the trading cost on returns, consider the random walk strategy that picks a random portfolio to randomly buy or sell: the mean economic rate of return drops from 0 percent with no trading costs ($b = 0$) to $-5.4$ percent with $b = 2$ basis points and $-13.5$ percent with $b = 5$ basis points. Returns drop by different amounts for different forecasting models since each models selects different option contracts into their portfolio and the price of these contracts (and thus the monetary effect of these basis points) varies across contracts. Thus, returns in the columns under the $b = 2$ header are not necessarily 5.4 percent lower than under the $b = 0$ header. In 2010-2012, returns in both euro and Canadian dollar options are still statistically significantly positive for $h = 1$ and possibly $h = 2$. For euro options in 2013-2015, they remain positive also for $h = 25$ and possibly $h = 20$; for Canadian dollar options in 2013-2015, returns are no longer positive. Even when negative, euro returns almost universally are significantly higher than those of the random walk strategy; for the Canadian dollar, this is true only in 2010-2012 for AVEL2 trades and for CFL10 and FXVIXC with $h \leq 5$. The story for $b = 5$ is virtually the same.[13]

These results are generally in line with the evidence about the quality of forecasting gathered in Section 5. Models outperform the random walk more often for $h = 1$ and $h = 2$. For euro 2010-2012, FXVIXL1, FXVIXL2, AVEL1 and AVEL2 models all do well, and their returns are better also. For euro 2013-2015, the performance of DI10FD for long horizons is surprising: forecasts are worse than the random walk model and AVEL2, yet returns are substantially better. With respect to the Canadian dollar in 2010-2012, the trading results correspond with the forecasting results. However, results for Canadian dollar options in 2013-2015 are the least expected: gains in forecasting performance do not turn into returns to trading, at least with the heuristic portfolio approach that derives option contract choice and buy/sell signals from the deviation between the forecasted and current implied volatility values.

A final comment compares the different factor models. CFL10 tends to perform worse than the AVE- and FXVIX-based models when $h$ rises and when transaction costs come into play. One may speculate that the variability of factor loadings in the CFL10 approach impart unwanted variability to forecasts. Aggregation weights in the cross-sectional average (AVE) or the market-wide volatility measure (FXVIX) are predetermined and do not vary with shocks in the IV data. Among FXVIX- and AVE-based models, AVEL2 tends to perform better than the other models that are specified in level variables; AVEC might have an edge over FXVIXC. Yet, Table 6 shows that FXVIX-based forecasts can sometimes be profitable as well.

## 7. Concluding remarks

This paper revisits the topic of forecasting the implied volatility surface of currency options with insights drawn from the spatial panel econometrics field. The concept of strong cross-sectional dependence draws the attention to underlying factors and variables that produce high correlations among observations located in "space." In the context of option implied volatility, this space is the implied volatility surface, define by time to maturity and moneyness. The familiar principal component analysis

[13] The rate of return declines virtually linearly with the rise in $b$. This is true even though the rate of return is a nonlinear function of $b$, since $b$ is small.

in the currency IVS forecasting literature is thus reinterpreted as generating the common factors that create strong cross-sectional dependence and, additionally, creating filtered IV variables that deserve to be forecasted as well. The unfiltered IVs thus consist of a mixture of time series processes, one defining the filter and another defining the filtered IVs. Moreover, the filter does not operate uniformly on the entire implied volatility surface but rather affects the different parts differently, raising one corner more than another as the filter variable rises. These effects are automatically built into the IVS forecasting process.

Furthermore, principal component factors are not the only way to generate common factors. The cross-sectional mean is another—and in this study proves to be a quite effective filter. Yet another type of filter may be found in observable variables that may reflect business cycles, environmental shocks, and the like. In this study, the FXVIX indicator that aggregates implied volatility across many currencies is a candidate and, similarly, proves to be quite effective.

We examine this class of models from various angles and find their forecasting performance more satisfactory than the existing models can offer. The random walk forecast can be beaten after all. Yet, when we examine the potential to profit from trading the implied volatility forecasts, we find that returns are greater than a random walk strategy of currency option trading but are not large enough to offset transaction costs.

In the light of the efficiency market hypothesis, it is not too surprising that trading implied volatility forecasts is profitable in the absence of transaction costs but is not profitable in the presence of transaction costs. In fact, this should be expected. There ought to be some useful information that can improve forecasts somewhat but not enough to cover the transaction cost of acting upon it.

**References**

Andrews, Donald W.K. (1991). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation." *Econometrica*, 59: 817–858.

Bai, Jushan, and Serena Ng (2002), "Determining the number of factors in approximate factor models." *Econometrica*, 70:191-221.

Bailey, Natalia, Sean Holly, and M. Hashem Pesaran (2016). "A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence." *Journal of Applied Econometrics,* 31: 249–280.

Binsbergen, Jules H., William F. Diamond, and Marco Grotteria (2019). "Risk-free interest rates." Working Paper 26138, National Bureau of Economic Research, Cambridge, MA.

Cao, Jie, and Bing Han (2013). "Cross section of option returns and idiosyncratic stock volatility." *Journal of Financial Economics*, 108: 231–249.

Cao, Jin, Bing Han, Xintong Zhan, Qing Tong (2021). "Option return predictability." *The Review of Financial Studies*, hhab067, https://doi.org/10.1093/rfs/hhab067.

Chudik, Alexander, M. Hashem Pesaran, and Elisa Tosetti (2011). "Weak and strong cross-section dependence and estimation of large panels." *Econometrics Journal*, 14: C45–C90.

Chalamandaris, Geogios, and Andrianos E. Tsekrekos (2010). "Predictable dynamics in implied volatility surfaces from OTC currency options." *Journal of Banking and Finance*, 34: 1175–1188.

Chalamandaris, Geogios, and Andrianos E. Tsekrekos (2011). "How important is the term structure in implied volatility surface modeling? Evidence from foreign exchange options." *Journal of International Money and Finance*, 30: 623–640.

Chalamandaris, Geogios, and Andrianos E. Tsekrekos (2014). "Predictability in implied volatility surfaces: Evidence from the euro OTC FX market." *European Journal of Finance*, 20(1): 33-58.

Christoffersen, Peter, and Kris Jacobs (2004). "The importance of the loss function in option valuation." *Journal of Financial Economics*, 72: 291–318.

Clark, Todd, and Michael McCracken (2013). "Advances in forecast evaluation." In G. Elliott and A. Timmermann, eds., *Handbook of Economic Forecasting*, Vol 2B, Ch. 20, 1107-1201, Elsevier B.V.

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

Diebold, Francis X., and Roberto S. Mariano (1995). "Comparing predictive accuracy." *Journal of Business & Economic Statistics*, 13(3): 253-263.

Dumas, Bernard, Jeff Fleming, and Robert Whaley (1998). "Implied volatility functions: Empirical tests." *Journal of Finance* 53, 2059–2106.

Elhorst, J. Paul, Jean-Loup Madre, and Alain Pirotte (2020). "Car traffic, habit persistence, cross-sectional dependence, and spatial heterogeneity: New insights using French departmental data." *Transportation Research Part A*, 132: 614-632.

Gibbons, Jean D., and Subhabrata Chakraborti (2003). *Nonparametric Statistical Inference*, 4$^{th}$ ed. Marcel Dekker Inc., New York.

Gonçalves, Sílvia, and Massimo Guidolin (2006). "Predictable dynamics in the S&P 500 index options implied volatility surface." *The Journal of Business*, 79(3): 1591-1635.

Guo, Biao, Qian Han, and Hai Lin (2017). "Are there gains from using information over the surface of implied volatilities?" *Journal of Futures Markets*, 38: 645-672.

Harvey, David I., Stephen J. Leybourne, and Paul Newbold (1998). "Tests for forecast encompassing." *Journal of Business & Economic Statistics*, 16(2): 254-259.

Heston, Steven, and Saikat Nandi (2000). "A closed-form GARCH option valuation model." *Review of Financial Studies*, 13: 585–625.

Hodrick, Robert J. (1992). "Dividend yields and expected stock returns: Alternative procedures for inference and measurement." *The Review of Financial Studies*, 5(3): 357-386.

Newey W.K., and K.D. West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica*, 55: 703–708.

Pesaran, M. Hashem (2015). "Testing weak cross-sectional dependence in large panels." *Econometric Reviews*, 34: 1089–1117.

Rapach, David E., and Mark E. Wohar (2007). "Forecasting the recent behavior of US business fixed investment spending: an analysis of competing models." *Journal of Forecasting*, 26: 33-51.

Rapach, David, and Guofu Zhou (2013). "Forecasting stock returns." In G. Elliott and A. Timmermann, Eds., *Handbook of Economic Forecasting*, Vol 2A, 329-383.

Stock, James, and Mark Watson (2002a). "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, December, 97(460), 1167-1179.

Stock, James, and Mark Watson (2002b). "Macroeconomic forecasting using diffusion indexes," *Journal of Business and Economic Statistics*, December, 20, 147-162.

Zeileis, Achim, Thomas Lumley, Nathaniel Graham, and Susanne Koell (2021). Package 'sandwich'. http://sandwich.R-Forge.R-project.org/.

**Appendix**

## A.1 Additional estimation results

### *A.1.1 Euro*

**Figure A.1**: Percentage difference in RMSPE relative to random walk, by forecast horizon h, delta and maturity: Euro, 2010-2012

A: ARIMA model

| | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -0.03 | 0.19 | -0.61 | -0.61 | 0.27 |
| | 2mo | -0.86 | -0.42 | -0.28 | -0.04 | 0.12 |
| | 3mo | -1.38 | -0.64 | 0.08 | -0.56 | -0.86 |
| | 6mo | -0.96 | -0.31 | -0.53 | -0.11 | -0.19 |
| | 12mo | -1.13 | -0.71 | -1.29 | -0.77 | -0.01 |
| h=2 | 1mo | 1.76 | 1.00 | 0.98 | -0.39 | 1.06 |
| | 2mo | -0.14 | 0.12 | 0.25 | 0.29 | 0.48 |
| | 3mo | -0.45 | 0.14 | 0.40 | 0.01 | 0.22 |
| | 6mo | -0.85 | -0.08 | 0.27 | 0.50 | 0.37 |
| | 12mo | -0.82 | -0.25 | -0.24 | -0.03 | 0.12 |
| h=5 | 1mo | 1.97 | 2.45 | 1.72 | -0.30 | -0.06 |
| | 2mo | -0.29 | -0.13 | -0.03 | 0.02 | -0.29 |
| | 3mo | -0.41 | -0.15 | -0.02 | -0.31 | -0.26 |
| | 6mo | -0.68 | -0.32 | -0.09 | 0.09 | 0.10 |
| | 12mo | -0.58 | -0.29 | -0.37 | -0.09 | -0.18 |
| h=10 | 1mo | 1.51 | 2.00 | 1.68 | 0.35 | 0.29 |
| | 2mo | -0.02 | -0.09 | -0.14 | -0.12 | -0.37 |
| | 3mo | -0.09 | -0.01 | -0.15 | -0.22 | -0.12 |
| | 6mo | -0.30 | -0.16 | 0.03 | 0.08 | 0.07 |
| | 12mo | -0.24 | -0.12 | -0.08 | 0.10 | 0.02 |
| h=15 | 1mo | 0.73 | 1.14 | 0.92 | 0.44 | 0.33 |
| | 2mo | -0.11 | -0.20 | -0.30 | -0.23 | -0.38 |
| | 3mo | -0.14 | -0.04 | -0.22 | 0.02 | -0.10 |
| | 6mo | -0.27 | -0.21 | -0.04 | 0.03 | 0.06 |
| | 12mo | -0.15 | -0.13 | -0.02 | 0.14 | 0.06 |
| h=20 | 1mo | 0.44 | 0.54 | 0.55 | 0.60 | 0.30 |
| | 2mo | -0.11 | -0.15 | -0.32 | -0.28 | -0.30 |
| | 3mo | -0.08 | -0.01 | -0.24 | 0.17 | -0.11 |
| | 6mo | -0.22 | -0.12 | -0.03 | 0.02 | 0.04 |
| | 12mo | -0.04 | 0.01 | 0.05 | 0.21 | 0.04 |
| h=25 | 1mo | -0.03 | -0.09 | 0.02 | 0.63 | 0.24 |
| | 2mo | -0.10 | -0.03 | -0.21 | -0.21 | -0.19 |
| | 3mo | -0.06 | -0.01 | -0.08 | 0.21 | -0.24 |
| | 6mo | -0.22 | -0.12 | -0.06 | 0.00 | 0.05 |
| | 12mo | -0.11 | -0.05 | 0.03 | 0.17 | 0.00 |
| | | 10 | 25 | atm | 75 | 90 |

B: AVEL2 model

| | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -1.09 | -1.58 | -2.41 | -2.41 | -1.39 |
| | 2mo | -1.48 | -0.90 | -1.41 | -0.70 | -1.03 |
| | 3mo | -1.33 | -1.22 | -1.48 | -0.35 | -0.55 |
| | 6mo | -1.09 | -0.20 | -0.73 | -0.66 | -0.90 |
| | 12mo | -1.47 | -0.88 | -0.80 | -0.57 | -0.90 |
| h=2 | 1mo | 0.20 | -0.49 | -2.76 | -3.26 | -2.10 |
| | 2mo | 0.65 | 0.87 | -0.05 | 0.19 | 0.35 |
| | 3mo | 0.00 | 0.13 | -0.48 | -0.42 | -0.14 |
| | 6mo | -1.03 | -0.38 | -0.33 | -0.32 | -0.14 |
| | 12mo | -1.48 | -1.20 | -0.51 | -0.36 | -0.17 |
| h=5 | 1mo | 0.36 | 0.84 | -2.21 | -3.35 | -3.07 |
| | 2mo | -0.07 | -0.04 | -1.90 | -1.44 | -0.22 |
| | 3mo | -0.18 | 0.19 | -0.19 | -0.60 | -0.71 |
| | 6mo | -2.46 | -1.39 | -0.28 | -0.08 | 0.27 |
| | 12mo | -3.36 | -2.84 | -0.71 | -0.03 | 0.05 |
| h=10 | 1mo | 0.54 | 1.17 | -4.52 | -6.87 | -6.69 |
| | 2mo | -0.17 | -0.27 | -2.89 | -2.79 | 0.05 |
| | 3mo | -0.16 | 0.83 | 0.05 | -0.31 | -0.71 |
| | 6mo | -4.71 | -2.27 | 0.40 | -0.34 | 0.18 |
| | 12mo | -5.54 | -4.58 | -1.06 | -0.34 | 0.18 |
| h=15 | 1mo | 0.52 | 0.89 | -5.80 | -8.96 | -9.08 |
| | 2mo | 0.03 | -0.47 | -3.62 | -3.91 | 0.60 |
| | 3mo | 0.04 | 1.56 | 0.38 | -0.28 | -0.77 |
| | 6mo | -5.50 | -2.50 | 0.93 | -0.48 | -0.02 |
| | 12mo | -7.16 | -5.10 | -1.45 | -0.60 | 0.01 |
| h=20 | 1mo | 0.71 | 1.12 | -6.64 | -10.22 | -10.52 |
| | 2mo | 0.18 | -0.62 | -4.17 | -4.73 | 0.79 |
| | 3mo | 0.20 | 1.92 | 0.22 | -0.13 | -1.04 |
| | 6mo | -5.40 | -2.76 | 0.91 | -0.37 | -0.06 |
| | 12mo | -8.16 | -5.36 | -2.28 | -0.87 | 0.06 |
| h=25 | 1mo | 1.08 | 1.59 | -6.64 | -10.16 | -10.86 |
| | 2mo | 0.47 | -0.53 | -4.39 | -4.90 | 1.41 |
| | 3mo | 0.36 | 2.34 | 0.19 | 0.60 | -0.90 |
| | 6mo | -4.45 | -2.68 | 1.15 | -0.05 | 0.31 |
| | 12mo | -7.63 | -4.85 | -2.29 | -0.63 | 0.75 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| | | | |
|---|---|---|---|
| 1% | 5% | 10% | 20% |

**Table A.1**: Aggregate out-of-sample R2 and MDM/MCW tests, and summary of separate MDM/MCW and sign tests: Euro 2013-2015

| Type of Model | Forecast horizon | | | | | | | Share preferred[a] | | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| | $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Model | R.Walk | |
| ARMA | -0.014 | -0.015 | -0.049 | -0.092 | -0.135 | -0.171 | -0.224 | 1.71 | n.a. | 0.57 |
| ARIMA | -0.003 * | -0.006 * | -0.003 | 0.002 * | 0.006 ** | 0.005 *** | 0.006 *** | 22.86 | n.a. | 5.71 |
| CFC3 | -0.009 | -0.021 + | -0.017 ++ | -0.018 | -0.023 | -0.039 + | -0.047 + | 0.00 | 25.71 | 0.00 |
| CFC3D | 0.018 | 0.008 | -0.004 | -0.005 | -0.012 | -0.026 | -0.032 | 2.29 | 18.29 | 9.71 |
| CFL3 | -0.156 +++ | -0.054 +++ | -0.013 | 0.005 | 0.011 | 0.007 | 0.009 | 16.00 | 41.71 | 20.00 |
| CFL10 | 0.007 | 0.009 | 0.008 | 0.012 | 0.014 | 0.009 | 0.011 | 22.29 | 20.57 | 22.29 |
| GG | 0.022 | 0.027 ** | 0.004 | -0.005 | -0.008 | -0.023 | -0.040 | 19.43 | 8.57 | 19.43 |
| DI5 | 0.002 | -0.001 | -0.004 | -0.002 | -0.010 | -0.014 | -0.023 | 0.00 | 0.00 | 0.57 |
| DI10 | 0.000 | -0.002 | -0.009 | -0.003 | -0.015 | -0.017 | -0.028 | 0.00 | 2.86 | 3.43 |
| DI5F | 0.002 | -0.019 + | -0.036 + | -0.028 + | -0.029 | -0.030 | -0.043 + | 0.00 | 29.14 | 0.00 |
| DI10F | 0.000 | -0.019 | -0.039 ++ | -0.029 ++ | -0.032 | -0.037 | -0.050 + | 0.00 | 30.29 | 1.71 |
| DI1FD | 0.015 | 0.004 | -0.009 | -0.017 | -0.023 | -0.037 | -0.046 + | 2.86 | 23.43 | 7.43 |
| DI5FD | 0.027 | -0.003 | -0.025 | -0.024 | -0.027 | -0.030 | -0.041 | 2.86 | 28.00 | 8.00 |
| DI10FD | 0.022 | -0.003 | -0.026 | -0.018 | -0.023 | -0.029 | -0.043 + | 1.71 | 29.14 | 5.71 |
| AR2p5 | 0.005 ** | 0.005 * | -0.001 | -0.009 | -0.017 | -0.027 | -0.036 | 10.29 | n.a. | 1.71 |
| AR2p5D | 0.027 *** | 0.016 *** | 0.002 | -0.001 | -0.012 | -0.019 | -0.029 | 16.57 | n.a. | 7.43 |
| SARpdP | 0.007 | 0.008 | 0.001 | -0.009 | -0.018 | -0.029 | -0.038 | 13.71 | n.a. | 1.14 |
| SAR211 | 0.005 | 0.006 | 0.000 | -0.008 | -0.016 | -0.026 | -0.035 | 10.29 | n.a. | 1.71 |
| FXVIXC | 0.016 *** | 0.001 ** | -0.006 | -0.014 | -0.022 | -0.033 | -0.042 | 12.57 | n.a. | 6.29 |
| AVEC | 0.028 *** | 0.015 *** | -0.002 | -0.003 | -0.012 | -0.019 | -0.029 | 14.86 | n.a. | 9.71 |
| FXVIXL1 | -0.046 | -0.068 | -0.069 | -0.080 | -0.053 | -0.021 | 0.014 | 28.57 | n.a. | 7.43 |
| FXVIXL2 | -0.042 * | -0.037 | -0.042 | -0.033 | -0.015 * | 0.002 * | 0.016 * | 29.71 | n.a. | 4.57 |
| AVEL1 | 0.019 | 0.018 | 0.008 | 0.011 | 0.015 | 0.016 | 0.021 | 41.71 | n.a. | 16.00 |
| AVEL2 | 0.015 | 0.014 | 0.005 | 0.005 | 0.008 | 0.008 | 0.014 | 40.00 | n.a. | 17.71 |

Notes: [a] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable. Significance levels: one-sided tests in favor of indicated model: *** 1%, ** 5%, * 10%; one-sided tests in favor of random walk model: +++ 1%, ++ 5%, + 10%.

**Figure A.2**: Percentage difference in RMSPE relative to random walk, by forecast horizon h, delta and maturity: Euro, 2013-2015

A: ARIMA model

| h | Maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -0.77 | -0.27 | 0.06 | 1.03 | 1.36 |
| | 2mo | 0.27 | 0.37 | 0.60 | 0.36 | 0.16 |
| | 3mo | -0.26 | 0.15 | -0.58 | -0.28 | -0.05 |
| | 6mo | -0.79 | 0.31 | -0.07 | 0.41 | 0.34 |
| | 12mo | -0.42 | -0.41 | -0.08 | -0.10 | -0.04 |
| h=2 | 1mo | -0.61 | -0.36 | -0.12 | 1.36 | 1.16 |
| | 2mo | 0.42 | 0.40 | 0.61 | 0.44 | 0.14 |
| | 3mo | 0.15 | 0.15 | -0.43 | -0.16 | 0.16 |
| | 6mo | -0.44 | 0.40 | 0.65 | 1.60 | 1.61 |
| | 12mo | 0.46 | 0.39 | 0.32 | 0.10 | 0.08 |
| h=5 | 1mo | 0.03 | 0.14 | 0.18 | 0.18 | 0.12 |
| | 2mo | 0.25 | 0.22 | 0.11 | 0.02 | -0.03 |
| | 3mo | 0.68 | 0.12 | -0.22 | -0.14 | 0.13 |
| | 6mo | 0.18 | 0.52 | 0.52 | 0.35 | 0.44 |
| | 12mo | 0.44 | 0.47 | 0.29 | 0.13 | 0.02 |
| h=10 | 1mo | -0.40 | -0.58 | -0.70 | -0.63 | -0.32 |
| | 2mo | 0.29 | 0.25 | 0.27 | 0.18 | 0.06 |
| | 3mo | 0.51 | -0.07 | -0.28 | -0.17 | 0.02 |
| | 6mo | 0.30 | 0.34 | 0.24 | 0.42 | 0.61 |
| | 12mo | 0.30 | 0.37 | 0.24 | 0.24 | 0.16 |
| h=15 | 1mo | -0.84 | -1.00 | -0.85 | -0.74 | -0.25 |
| | 2mo | 0.18 | 0.16 | 0.11 | 0.05 | -0.01 |
| | 3mo | 0.09 | -0.23 | -0.40 | -0.41 | -0.34 |
| | 6mo | 0.07 | 0.20 | 0.02 | -0.13 | 0.02 |
| | 12mo | 0.14 | 0.17 | 0.38 | 0.16 | 0.09 |
| h=20 | 1mo | -0.37 | -0.64 | -0.76 | -0.89 | -0.34 |
| | 2mo | 0.15 | 0.16 | 0.08 | 0.08 | 0.07 |
| | 3mo | 0.11 | -0.21 | -0.40 | -0.46 | -0.39 |
| | 6mo | 0.11 | 0.21 | -0.05 | -0.19 | -0.02 |
| | 12mo | 0.03 | -0.03 | 0.12 | 0.19 | 0.13 |
| h=25 | 1mo | -0.42 | -1.01 | -0.86 | -0.78 | -0.37 |
| | 2mo | 0.17 | 0.18 | 0.10 | 0.08 | 0.08 |
| | 3mo | 0.12 | -0.21 | -0.50 | -0.63 | -0.62 |
| | 6mo | 0.18 | 0.26 | 0.00 | -0.31 | -0.18 |
| | 12mo | 0.10 | -0.01 | 0.29 | 0.25 | 0.18 |
| | | 10 | 25 | atm | 75 | 90 |

B: AVEC model

| h | Maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -1.81 | -1.78 | -1.85 | -1.77 | -1.71 |
| | 2mo | -2.53 | -2.55 | -2.66 | -2.66 | -2.69 |
| | 3mo | -1.13 | -0.98 | -1.00 | -0.86 | -0.80 |
| | 6mo | -0.28 | 0.06 | 0.17 | 0.43 | 0.35 |
| | 12mo | 0.07 | 0.06 | 0.58 | 0.66 | 0.88 |
| h=2 | 1mo | -1.80 | -2.01 | -2.33 | -2.46 | -2.52 |
| | 2mo | -0.57 | -0.72 | -0.94 | -1.01 | -1.02 |
| | 3mo | 0.89 | 0.80 | 0.65 | 0.53 | 0.51 |
| | 6mo | 0.98 | 0.97 | 0.78 | 0.75 | 0.66 |
| | 12mo | 1.65 | 1.41 | 1.26 | 0.99 | 0.88 |
| h=5 | 1mo | -0.22 | -0.43 | -0.45 | -0.33 | -0.15 |
| | 2mo | 0.53 | 0.29 | 0.06 | 0.03 | 0.05 |
| | 3mo | 0.91 | 0.56 | 0.23 | 0.06 | 0.15 |
| | 6mo | 1.18 | 0.78 | 0.24 | -0.01 | -0.06 |
| | 12mo | 1.42 | 0.96 | 0.41 | 0.01 | -0.14 |
| h=10 | 1mo | 0.16 | -0.02 | -0.05 | 0.02 | 0.18 |
| | 2mo | 0.23 | 0.00 | -0.20 | -0.20 | -0.10 |
| | 3mo | 0.61 | 0.31 | 0.10 | 0.03 | 0.18 |
| | 6mo | 0.81 | 0.50 | 0.11 | -0.04 | 0.04 |
| | 12mo | 1.07 | 0.74 | 0.37 | 0.12 | 0.11 |
| h=15 | 1mo | 0.63 | 0.51 | 0.49 | 0.52 | 0.61 |
| | 2mo | 0.60 | 0.42 | 0.28 | 0.31 | 0.45 |
| | 3mo | 0.95 | 0.72 | 0.55 | 0.50 | 0.60 |
| | 6mo | 1.18 | 0.92 | 0.63 | 0.54 | 0.65 |
| | 12mo | 1.40 | 1.15 | 0.91 | 0.80 | 0.82 |
| h=20 | 1mo | 1.20 | 1.05 | 0.93 | 0.83 | 0.83 |
| | 2mo | 1.04 | 0.82 | 0.62 | 0.55 | 0.61 |
| | 3mo | 1.35 | 1.10 | 0.87 | 0.75 | 0.79 |
| | 6mo | 1.51 | 1.22 | 0.88 | 0.77 | 0.86 |
| | 12mo | 1.70 | 1.39 | 1.12 | 1.02 | 1.09 |
| h=25 | 1mo | 1.68 | 1.55 | 1.46 | 1.37 | 1.36 |
| | 2mo | 1.58 | 1.35 | 1.14 | 1.03 | 1.06 |
| | 3mo | 1.75 | 1.48 | 1.22 | 1.08 | 1.12 |
| | 6mo | 1.92 | 1.58 | 1.23 | 1.11 | 1.27 |
| | 12mo | 2.17 | 1.81 | 1.49 | 1.37 | 1.55 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| | | | |
|---|---|---|---|
| 1% | 5% | 10% | 20% |

**Figure A.2**: Continued

C: CFL10 model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -1.90 | -1.65 | -0.75 | -0.84 | -1.04 |
| | 2mo | -0.40 | -0.25 | 0.43 | 0.59 | 0.62 |
| | 3mo | -0.83 | -0.68 | 0.76 | 1.25 | 1.10 |
| | 6mo | -0.19 | 0.07 | 0.87 | 0.78 | 0.71 |
| | 12mo | -1.53 | 0.11 | 1.28 | 0.95 | -0.50 |
| h=2 | 1mo | -2.52 | -2.22 | -1.85 | -2.03 | -2.12 |
| | 2mo | 0.65 | 0.88 | 1.37 | 1.18 | 1.31 |
| | 3mo | -0.19 | -0.04 | 1.09 | 1.43 | 2.01 |
| | 6mo | -0.12 | -0.08 | 0.62 | 0.62 | 0.97 |
| | 12mo | -1.52 | -0.56 | 0.33 | 0.10 | -0.47 |
| h=5 | 1mo | -3.15 | -2.74 | -2.68 | -2.89 | -2.85 |
| | 2mo | 2.95 | 2.73 | 2.40 | 1.48 | 1.37 |
| | 3mo | 1.81 | 1.61 | 1.79 | 1.32 | 1.86 |
| | 6mo | 1.18 | 0.91 | 0.80 | 0.52 | 1.04 |
| | 12mo | -2.59 | -2.41 | -2.65 | -3.09 | -2.16 |
| h=10 | 1mo | -3.56 | -3.19 | -3.41 | -3.99 | -4.33 |
| | 2mo | 3.65 | 3.26 | 2.84 | 1.71 | 1.11 |
| | 3mo | 2.90 | 2.42 | 1.66 | 0.62 | 0.11 |
| | 6mo | 2.60 | 2.15 | 1.36 | 0.14 | -0.70 |
| | 12mo | -2.72 | -2.79 | -3.68 | -4.89 | -4.84 |
| h=15 | 1mo | -2.77 | -2.66 | -3.35 | -4.37 | -5.05 |
| | 2mo | 5.05 | 4.43 | 3.50 | 1.56 | 0.17 |
| | 3mo | 3.99 | 3.12 | 1.64 | -0.27 | -1.62 |
| | 6mo | 3.11 | 2.52 | 0.92 | -1.50 | -3.95 |
| | 12mo | -3.42 | -3.39 | -4.45 | -6.42 | -7.51 |
| h=20 | 1mo | -1.98 | -2.15 | -2.97 | -4.12 | -4.90 |
| | 2mo | 5.34 | 4.86 | 3.92 | 1.98 | 0.24 |
| | 3mo | 4.57 | 3.80 | 2.32 | 0.19 | -1.60 |
| | 6mo | 2.84 | 2.61 | 1.10 | -1.67 | -5.02 |
| | 12mo | -4.41 | -4.00 | -4.82 | -6.83 | -8.88 |
| h=25 | 1mo | -1.96 | -2.16 | -2.95 | -4.21 | -4.94 |
| | 2mo | 4.47 | 4.14 | 3.43 | 1.74 | 0.07 |
| | 3mo | 4.34 | 4.01 | 2.84 | 0.66 | -1.49 |
| | 6mo | 1.81 | 2.17 | 1.46 | -1.14 | -4.91 |
| | 12mo | -5.93 | -4.97 | -4.74 | -6.13 | -8.78 |
| | | 10 | 25 | atm | 75 | 90 |

D: DI10FD model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -1.66 | -1.65 | -1.85 | -1.87 | -1.83 |
| | 2mo | -2.36 | -2.29 | -1.93 | -1.58 | -1.62 |
| | 3mo | -1.14 | -1.00 | -0.70 | -0.16 | -0.14 |
| | 6mo | -0.09 | 0.25 | 0.70 | 1.27 | 1.52 |
| | 12mo | 0.48 | 0.60 | 1.03 | 1.07 | 1.34 |
| h=2 | 1mo | -1.39 | -1.73 | -2.29 | -2.53 | -2.71 |
| | 2mo | 1.45 | 1.63 | 1.48 | 0.72 | 0.23 |
| | 3mo | 1.78 | 2.19 | 2.34 | 2.03 | 1.82 |
| | 6mo | 1.49 | 1.97 | 2.27 | 2.14 | 1.80 |
| | 12mo | 1.86 | 1.92 | 1.80 | 1.56 | 1.53 |
| h=5 | 1mo | 0.56 | 0.49 | 0.74 | 1.32 | 1.66 |
| | 2mo | 1.77 | 1.98 | 1.95 | 1.70 | 1.52 |
| | 3mo | 1.32 | 1.39 | 1.57 | 1.59 | 1.61 |
| | 6mo | 1.09 | 1.19 | 1.35 | 1.46 | 1.55 |
| | 12mo | 1.56 | 1.60 | 1.31 | 0.98 | 1.00 |
| h=10 | 1mo | -0.10 | -0.25 | -0.25 | -0.12 | 0.11 |
| | 2mo | 1.74 | 1.95 | 1.92 | 1.45 | 1.18 |
| | 3mo | 1.28 | 1.46 | 1.67 | 1.53 | 1.58 |
| | 6mo | 1.08 | 1.29 | 1.68 | 1.66 | 1.81 |
| | 12mo | 1.11 | 0.91 | 0.67 | 0.75 | 1.30 |
| h=15 | 1mo | 0.71 | 0.67 | 0.64 | 0.66 | 0.72 |
| | 2mo | 1.87 | 1.95 | 1.80 | 1.38 | 1.24 |
| | 3mo | 1.61 | 1.65 | 1.59 | 1.31 | 1.26 |
| | 6mo | 1.30 | 1.30 | 1.33 | 1.20 | 1.31 |
| | 12mo | 1.40 | 1.23 | 0.85 | 0.66 | 0.89 |
| h=20 | 1mo | 1.16 | 0.91 | 0.64 | 0.47 | 0.45 |
| | 2mo | 2.58 | 2.58 | 2.27 | 1.61 | 1.31 |
| | 3mo | 2.17 | 2.11 | 1.92 | 1.51 | 1.41 |
| | 6mo | 2.09 | 1.98 | 1.78 | 1.46 | 1.42 |
| | 12mo | 2.03 | 1.70 | 1.06 | 0.71 | 0.82 |
| h=25 | 1mo | 2.14 | 1.98 | 1.8 | 1.62 | 1.56 |
| | 2mo | 3.08 | 3.08 | 2.82 | 2.16 | 1.81 |
| | 3mo | 2.55 | 2.49 | 2.35 | 1.93 | 1.79 |
| | 6mo | 2.42 | 2.27 | 2.12 | 1.81 | 1.77 |
| | 12mo | 2.49 | 2.16 | 1.55 | 1.17 | 1.32 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance is coded by color and font thickness and italics:

| | | | | |
|---|---|---|---|---|
| In favor of indicated model: | 1% | 5% | 10% | 20% |
| In favor of random walk model: | 1% | 5% | 10% | 20% |

**Figure A.2**: Continued

E: FXVIX1 model

| | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | 2.08 | 2.46 | 3.17 | 3.68 | 3.82 |
| | 2mo | 1.75 | 1.92 | 2.26 | 2.29 | 1.97 |
| | 3mo | 1.28 | 1.43 | 1.48 | 1.56 | 1.46 |
| | 6mo | 1.69 | 1.88 | 2.56 | 2.75 | 2.31 |
| | 12mo | 0.70 | 0.97 | 1.56 | 2.23 | 1.80 |
| h=2 | 1mo | 2.76 | 3.16 | 3.92 | 4.46 | 4.72 |
| | 2mo | 2.94 | 3.36 | 4.11 | 3.99 | 3.74 |
| | 3mo | 2.56 | 2.89 | 2.92 | 3.05 | 3.48 |
| | 6mo | 2.74 | 2.73 | 3.34 | 3.30 | 2.77 |
| | 12mo | 0.88 | 1.20 | 1.83 | 2.36 | 2.09 |
| h=5 | 1mo | 3.41 | 3.85 | 4.64 | 5.27 | 5.67 |
| | 2mo | 2.16 | 2.82 | 4.03 | 4.01 | 3.94 |
| | 3mo | 1.66 | 2.50 | 3.02 | 3.44 | 3.94 |
| | 6mo | 2.46 | 1.68 | 2.99 | 3.18 | 2.62 |
| | 12mo | **-0.79** | -0.32 | 0.52 | 0.97 | 0.83 |
| h=10 | 1mo | 5.31 | 5.83 | 6.71 | 7.33 | 7.78 |
| | 2mo | 2.83 | 2.90 | 4.16 | 4.10 | 3.89 |
| | 3mo | 1.90 | 2.32 | 1.94 | 2.08 | 2.61 |
| | 6mo | 3.85 | 1.13 | 2.87 | 2.33 | 1.69 |
| | 12mo | **-2.16** | **-1.78** | -0.90 | -0.82 | -1.46 |
| h=15 | 1mo | 4.47 | 4.63 | 5.15 | **5.60** | **6.10** |
| | 2mo | 3.43 | 2.42 | 2.91 | 2.54 | 2.29 |
| | 3mo | 2.40 | 2.39 | 0.75 | **0.19** | **0.95** |
| | 6mo | 4.71 | -0.52 | 1.06 | 0.19 | -0.39 |
| | 12mo | -4.98 | -4.67 | -3.64 | -4.09 | -4.77 |
| h=20 | 1mo | 1.52 | **1.28** | **1.76** | **2.55** | **3.80** |
| | 2mo | 3.21 | 1.26 | 1.57 | **1.15** | **1.25** |
| | 3mo | 3.10 | 2.53 | -0.14 | **-0.68** | **0.00** |
| | 6mo | 5.43 | -1.04 | **0.14** | -0.61 | -1.33 |
| | 12mo | **-7.17** | **-6.49** | -5.03 | -5.60 | -6.39 |
| h=25 | 1mo | **-1.72** | **-2.01** | **-1.25** | **-0.02** | **2.13** |
| | 2mo | 1.04 | -0.86 | 0.07 | -0.11 | 0.89 |
| | 3mo | 2.39 | 1.65 | -0.88 | **-1.17** | **-0.76** |
| | 6mo | 4.39 | -2.07 | **-1.03** | **-1.09** | **-2.29** |
| | 12mo | -9.73 | -8.79 | -6.01 | -5.11 | -5.84 |
| | | 10 | 25 | atm | 75 | 90 |

F: AVEL2 model

| | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -2.17 | -2.12 | -1.61 | -0.98 | -0.32 |
| | 2mo | -1.43 | -1.23 | **0** | 0.84 | 0.81 |
| | 3mo | -1.28 | -0.89 | 0.1 | 1.15 | 0.64 |
| | 6mo | -1.23 | -0.92 | **-0.32** | 0.66 | 0.48 |
| | 12mo | -1.48 | -1.07 | **-0.82** | **-0.55** | **-0.31** |
| h=2 | 1mo | -2.32 | -2.53 | -2.24 | -1.72 | -1.1 |
| | 2mo | -0.4 | -0.18 | 1.25 | 1.59 | 1.19 |
| | 3mo | **-0.51** | **-0.51** | 0.72 | 1.39 | 0.66 |
| | 6mo | **-0.56** | -1.02 | -0.62 | 0.37 | 0.28 |
| | 12mo | -0.76 | -0.8 | -0.83 | -0.97 | **-0.75** |
| h=5 | 1mo | **-1.43** | **-1.73** | -2.1 | -2.08 | **-1.57** |
| | 2mo | 1.65 | 1.46 | 1.99 | 1.5 | 0.83 |
| | 3mo | 1.28 | 0.96 | 1.27 | 0.98 | 0.38 |
| | 6mo | 1.16 | 0.25 | -0.4 | 0.03 | 0.38 |
| | 12mo | -0.52 | -0.62 | **-1.29** | **-2.37** | **-2.37** |
| h=10 | 1mo | **-1.52** | **-2.09** | **-2.87** | -3.31 | -3.1 |
| | 2mo | 3.39 | 2.72 | 2.75 | 1.76 | 0.56 |
| | 3mo | 2.79 | 2.06 | 1.66 | 0.86 | 0.18 |
| | 6mo | 0.85 | 1.31 | 0.27 | -0.35 | -0.42 |
| | 12mo | -0.85 | -0.53 | -1.97 | -3.82 | -4.66 |
| h=15 | 1mo | -0.47 | -1.47 | -2.88 | **-4.07** | -4.36 |
| | 2mo | 5.34 | 4.08 | 3.14 | 1.24 | -0.22 |
| | 3mo | 4.05 | 2.83 | 1.39 | -0.25 | -0.13 |
| | 6mo | -0.59 | 1.51 | -0.39 | -2.53 | -1.87 |
| | 12mo | -1.77 | -0.43 | -3.02 | -6.36 | -8.17 |
| h=20 | 1mo | -1.08 | -2.2 | **-3.54** | -4.58 | -4.66 |
| | 2mo | 5.56 | 4.37 | 3.27 | 1.53 | 0.21 |
| | 3mo | 4.54 | 3.35 | 1.93 | 0.24 | 0.19 |
| | 6mo | -0.82 | 1.66 | 0.28 | -1.97 | -1.34 |
| | 12mo | -2.17 | 0.09 | -3.47 | -6.68 | -8.97 |
| h=25 | 1mo | -2.02 | -3.13 | -4.44 | -5.5 | -5.33 |
| | 2mo | 4.71 | 3.53 | 2.48 | 1.15 | 0.29 |
| | 3mo | 4.56 | 3.58 | 2.31 | 0.7 | -0.53 |
| | 6mo | -0.77 | 1.65 | 1.28 | -0.62 | -0.6 |
| | 12mo | **-1.53** | 0.8 | -3.54 | -5.78 | -8.3 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| 1% | 5% | 10% | 20% |
|---|---|---|---|

### *A.1.2 Canadian Dollar*

**Table A.2**: Aggregate out-of-sample R2 and MDM/MCW tests, and summary of separate MDM/MCW and sign tests: Canadian Dollar 2010-2012

| Type of Model | Forecast horizon $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Share preferred[a] Model | R.Walk | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| ARMA | -0.039 | -0.081 | -0.205 | -0.340 | -0.411 | -0.450 | -0.467 | 2.86 | n.a. | 0.00 |
| ARIMA | -0.006 | -0.010 | -0.003 | -0.004 | -0.004 | -0.005 | -0.005 | 33.71 | n.a. | 22.86 |
| CFC3 | 0.004 | 0.002 | 0.017 | 0.024 | 0.027 | 0.031 | 0.036 | 6.86 | 0.00 | 69.14 |
| CFC3D | 0.037 | 0.045 * | 0.017 | 0.024 | 0.027 | 0.031 | 0.036 | 11.43 | 0.00 | 72.00 |
| CFL3 | -0.341 +++ | -0.159 +++ | -0.063 +++ | -0.034 ++ | -0.024 ++ | -0.017 | -0.013 | 4.57 | 31.43 | 9.14 |
| CFL10 | 0.011 | 0.005 | 0.000 | -0.008 | -0.009 | -0.008 | -0.007 | 4.57 | 1.71 | 5.14 |
| GG | 0.019 | 0.005 | -0.001 | -0.026 | -0.045 | -0.049 | -0.052 | 8.57 | 8.57 | 35.43 |
| DI5 | -0.016 | -0.027 | -0.027 | -0.007 | -0.001 | 0.009 | 0.022 | 1.14 | 2.29 | 57.71 |
| DI10 | -0.010 | -0.024 | -0.037 | -0.012 | -0.004 | 0.007 | 0.021 | 1.71 | 0.57 | 49.14 |
| DI5F | -0.010 | -0.045 | -0.020 | 0.011 | 0.009 | 0.017 | 0.030 | 1.14 | 4.00 | 53.71 |
| DI10F | -0.005 | -0.053 + | -0.023 | 0.004 | 0.004 | 0.016 | 0.031 | 1.14 | 6.29 | 48.00 |
| DI1FD | 0.038 | 0.045 ** | 0.017 | 0.024 | 0.027 | 0.031 | 0.036 | 10.86 | 0.00 | 72.00 |
| DI5FD | 0.025 | -0.006 | -0.008 | 0.022 | 0.013 | 0.020 | 0.032 | 4.57 | 1.14 | 57.71 |
| DI10FD | 0.022 | -0.018 | -0.007 | 0.016 | 0.009 | 0.020 | 0.033 | 4.57 | 2.29 | 57.71 |
| AR2p5 | -0.006 * | -0.004 | 0.011 | 0.015 | 0.014 | 0.023 | 0.029 | 17.14 | n.a. | 69.14 |
| AR2p5D | 0.024 *** | 0.033 *** | 0.000 | 0.007 | 0.009 | 0.018 | 0.028 | 26.29 | n.a. | 70.29 |
| SARpdP | -0.004 | -0.006 | 0.013 | 0.016 | 0.016 | 0.023 | 0.027 | 17.71 | n.a. | 73.14 |
| SAR211 | -0.005 | -0.004 | 0.012 | 0.015 | 0.014 | 0.021 | 0.026 | 15.43 | n.a. | 68.57 |
| FXVIXC | 0.058 *** | 0.061 *** | 0.025 ** | 0.025 | 0.024 | 0.028 | 0.035 | 52.57 | n.a. | 68.57 |
| AVEC | 0.034 *** | 0.037 *** | -0.004 | 0.005 | 0.008 | 0.017 | 0.027 | 30.29 | n.a. | 69.71 |
| FXVIXL1 | 0.019 | 0.035 | 0.023 | 0.033 | 0.046 | 0.071 | 0.095 | 60.57 | n.a. | 61.14 |
| FXVIXL2 | -0.008 ** | -0.028 * | -0.037 | -0.021 | 0.003 | 0.030 | 0.054 | 15.43 | n.a. | 5.71 |
| AVEL1 | 0.014 | 0.010 | 0.015 | 0.015 | 0.015 | 0.014 | 0.014 | 48.57 | n.a. | 31.43 |
| AVEL2 | 0.012 | 0.009 | 0.011 | 0.013 | 0.014 | 0.013 | 0.011 | 43.43 | n.a. | 27.43 |

Notes: [a] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable. Significance levels: one-sided tests in favor of indicated model: *** 1%, ** 5%, * 10%; one-sided tests in favor of random walk model: +++ 1%, ++ 5%, + 10%.

**Figure A.3**: Percentage difference in RMSPE relative to random walk, by forecast horizon h, delta and maturity: Canadian dollar, 2010-2012

A: ARIMA model

| h | Maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | 1.59 | 1.64 | 0.23 | -0.18 | -0.10 |
| | 2mo | 1.53 | 1.19 | 0.12 | **-0.46** | **-0.71** |
| | 3mo | 2.07 | 0.78 | **-0.19** | **-1.04** | **-1.54** |
| | 6mo | 1.71 | 1.26 | 0.10 | **-1.19** | **-1.15** |
| | 12mo | 0.90 | 0.28 | **-0.51** | **-0.39** | **-0.46** |
| h=2 | 1mo | 1.86 | 1.16 | 0.19 | 0.07 | 0.07 |
| | 2mo | 1.59 | 1.69 | 0.19 | -0.06 | -0.29 |
| | 3mo | 1.66 | 1.14 | 0.06 | -0.30 | **-0.52** |
| | 6mo | 1.41 | 1.22 | 0.59 | -0.35 | **-0.44** |
| | 12mo | 1.04 | 0.49 | -0.15 | 0.19 | 0.07 |
| h=5 | 1mo | 2.21 | 2.25 | **-0.40** | **-0.46** | -0.26 |
| | 2mo | 1.88 | 1.60 | **-0.39** | **-0.54** | **-0.50** |
| | 3mo | 1.50 | 1.69 | -0.35 | -0.52 | **-0.57** |
| | 6mo | -0.06 | -0.30 | **-0.63** | **-1.15** | **-1.11** |
| | 12mo | 0.29 | -0.20 | **-0.71** | 0.19 | **-0.66** |
| h=10 | 1mo | 2.72 | 2.62 | **-0.22** | -0.23 | -0.17 |
| | 2mo | 1.25 | 1.80 | -0.28 | -0.39 | -0.42 |
| | 3mo | 1.84 | 1.84 | -0.38 | -0.52 | -0.57 |
| | 6mo | -0.28 | -0.46 | **-0.62** | **-0.85** | **-0.88** |
| | 12mo | -0.15 | -0.52 | **-0.76** | -0.41 | **-0.77** |
| h=15 | 1mo | 3.11 | 2.15 | **-0.22** | **-0.26** | **-0.22** |
| | 2mo | 1.24 | 1.53 | **-0.27** | **-0.32** | **-0.34** |
| | 3mo | 1.82 | 1.78 | -0.20 | **-0.39** | **-0.48** |
| | 6mo | **-0.30** | **-0.37** | **-0.44** | **-0.61** | **-0.71** |
| | 12mo | -0.20 | -0.39 | **-0.49** | **-0.53** | **-0.64** |
| h=20 | 1mo | 2.06 | 1.94 | -0.03 | -0.14 | -0.17 |
| | 2mo | 1.14 | 1.52 | -0.16 | **-0.24** | **-0.27** |
| | 3mo | 1.64 | 1.58 | -0.07 | **-0.31** | **-0.38** |
| | 6mo | -0.11 | -0.19 | **-0.30** | **-0.42** | **-0.48** |
| | 12mo | -0.09 | -0.22 | **-0.30** | **-0.45** | **-0.40** |
| h=25 | 1mo | 1.78 | 1.63 | 0.03 | -0.03 | -0.07 |
| | 2mo | 1.10 | 1.43 | -0.09 | -0.16 | -0.14 |
| | 3mo | 1.66 | 1.64 | -0.01 | -0.19 | -0.19 |
| | 6mo | -0.18 | -0.21 | -0.25 | **-0.31** | **-0.31** |
| | 12mo | -0.15 | -0.26 | -0.30 | **-0.39** | **-0.26** |
| | | 10 | 25 | atm | 75 | 90 |

B: AVEC model

| h | Maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | **-6.73** | **-6.89** | **-6.24** | **-5.46** | **-4.11** |
| | 2mo | 1.18 | **1.28** | **1.07** | **0.93** | **1.07** |
| | 3mo | **0.70** | **0.75** | **0.55** | **0.23** | **0.12** |
| | 6mo | **-0.66** | **0.80** | **0.59** | **0.05** | **-0.18** |
| | 12mo | **-1.98** | **0.65** | **0.47** | **-0.16** | **-0.26** |
| h=2 | 1mo | **-6.40** | **-6.51** | **-5.53** | **-4.86** | **-4.02** |
| | 2mo | **-0.27** | **-0.21** | **-0.23** | **-0.26** | **-0.34** |
| | 3mo | **0.20** | **0.02** | **0.05** | **-0.22** | **-0.36** |
| | 6mo | 0.53 | 0.85 | 0.77 | **0.26** | **-0.17** |
| | 12mo | **-1.34** | **0.01** | 0.38 | **-0.25** | **0.05** |
| h=5 | 1mo | 0.04 | 0.01 | 0.12 | 0.14 | 0.09 |
| | 2mo | 0.30 | 0.32 | 0.44 | 0.49 | 0.47 |
| | 3mo | -0.20 | 0.04 | 0.27 | 0.38 | 0.38 |
| | 6mo | -0.07 | 0.31 | 0.47 | 0.47 | 0.43 |
| | 12mo | -1.11 | -0.43 | 0.11 | 0.34 | 0.73 |
| h=10 | 1mo | -0.38 | -0.23 | -0.05 | 0.11 | 0.27 |
| | 2mo | -0.58 | -0.39 | -0.12 | 0.10 | 0.28 |
| | 3mo | -0.89 | -0.60 | -0.27 | 0.00 | 0.18 |
| | 6mo | -1.45 | -1.07 | -0.64 | -0.30 | -0.06 |
| | 12mo | **-2.48** | -1.91 | -1.08 | -0.56 | 0.02 |
| h=15 | 1mo | -0.51 | -0.29 | 0.03 | 0.27 | 0.52 |
| | 2mo | -0.97 | -0.70 | -0.24 | 0.14 | 0.50 |
| | 3mo | -1.47 | -1.08 | -0.47 | 0.01 | 0.38 |
| | 6mo | -2.54 | -2.00 | -1.11 | -0.43 | 0.03 |
| | 12mo | **-3.76** | **-3.23** | -1.73 | -0.85 | -0.04 |
| h=20 | 1mo | -1.07 | -0.78 | -0.35 | -0.01 | 0.34 |
| | 2mo | -1.54 | -1.26 | -0.66 | -0.17 | 0.31 |
| | 3mo | -2.21 | -1.80 | -1.02 | -0.36 | 0.16 |
| | 6mo | -3.73 | -3.18 | -1.97 | -0.98 | -0.28 |
| | 12mo | **-5.29** | **-4.84** | -2.89 | -1.49 | -0.41 |
| h=25 | 1mo | -1.46 | -1.13 | -0.65 | -0.27 | 0.13 |
| | 2mo | -2.09 | -1.74 | -1.04 | -0.47 | 0.06 |
| | 3mo | -2.88 | -2.41 | -1.52 | -0.74 | -0.13 |
| | 6mo | **-4.78** | -4.12 | -2.72 | -1.51 | -0.65 |
| | 12mo | -6.97 | **-6.29** | -4.00 | -2.26 | -1.01 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| | | | |
|---|---|---|---|
| **1%** | **5%** | **10%** | 20% |

**Figure A.3**: Continued

C: CFL10 model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -0.91 | -2.28 | -1.69 | -1.37 | -1.97 |
| | 2mo | 0.90 | 1.43 | 1.25 | 1.21 | -0.12 |
| | 3mo | 1.47 | 1.35 | 0.14 | -1.27 | -1.97 |
| | 6mo | 0.81 | 1.20 | -0.95 | -1.53 | -3.00 |
| | 12mo | -4.15 | 4.11 | 1.07 | 1.09 | -4.38 |
| h=2 | 1mo | -1.37 | -1.78 | -1.09 | -0.36 | -0.58 |
| | 2mo | 0.43 | 1.51 | 1.32 | 1.22 | -0.16 |
| | 3mo | 1.05 | 0.44 | -0.20 | -0.88 | -1.61 |
| | 6mo | 3.37 | 0.91 | -1.02 | -0.91 | -1.82 |
| | 12mo | -1.03 | 2.49 | 0.20 | 0.20 | -2.70 |
| h=5 | 1mo | 0.31 | 0.68 | 1.11 | 1.89 | 1.27 |
| | 2mo | -0.30 | 0.72 | 0.55 | 0.59 | -0.89 |
| | 3mo | -1.65 | -0.50 | -0.88 | -0.83 | -1.38 |
| | 6mo | 1.19 | -0.35 | -1.87 | -1.54 | -1.19 |
| | 12mo | 0.74 | 1.54 | -1.65 | -1.67 | -2.55 |
| h=10 | 1mo | 1.04 | 2.25 | 2.57 | 3.33 | 2.82 |
| | 2mo | -0.01 | 1.59 | 1.08 | 1.07 | -0.51 |
| | 3mo | -1.92 | -0.27 | -0.91 | -0.69 | -1.07 |
| | 6mo | 0.47 | -0.87 | -2.75 | -2.21 | -1.27 |
| | 12mo | 2.10 | 0.78 | -3.53 | -3.04 | -2.30 |
| h=15 | 1mo | 1.27 | 2.58 | 3.02 | 3.58 | 3.25 |
| | 2mo | -0.88 | 1.16 | 0.86 | 0.95 | -0.48 |
| | 3mo | -1.87 | -0.50 | -1.26 | -1.08 | -1.26 |
| | 6mo | 0.50 | -0.34 | -2.48 | -2.18 | -1.17 |
| | 12mo | 2.37 | 2.02 | -2.72 | -2.33 | -1.19 |
| h=20 | 1mo | 0.35 | 1.99 | 2.61 | 3.34 | 3.00 |
| | 2mo | -1.87 | 0.48 | 0.59 | 0.95 | -0.23 |
| | 3mo | -2.29 | -0.54 | -1.11 | -0.94 | -1.22 |
| | 6mo | 0.36 | 0.27 | -1.58 | -1.56 | -1.00 |
| | 12mo | 2.57 | 3.02 | -1.23 | -1.53 | -0.72 |
| h=25 | 1mo | -0.44 | 1.26 | 2.05 | 2.85 | 2.53 |
| | 2mo | -2.50 | -0.03 | 0.33 | 0.82 | -0.08 |
| | 3mo | -2.33 | -0.63 | -0.85 | -0.65 | -0.98 |
| | 6mo | 0.63 | 0.75 | -0.79 | -0.89 | -0.53 |
| | 12mo | 3.15 | 3.44 | -0.10 | -0.56 | 0.10 |
| | | 10 | 25 | atm | 75 | 90 |

D: DI10FD model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -5.98 | -5.76 | -4.52 | -4.13 | -3.63 |
| | 2mo | 0.88 | 1.47 | 1.48 | 1.13 | 0.76 |
| | 3mo | 1.44 | 1.85 | 1.58 | 0.76 | 0.97 |
| | 6mo | -4.04 | 0.83 | 1.58 | 0.63 | 0.65 |
| | 12mo | -2.45 | 1.68 | 2.25 | 1.24 | 1.74 |
| h=2 | 1mo | -1.81 | -1.69 | -0.97 | -1.33 | -1.44 |
| | 2mo | 3.85 | 4.90 | 5.53 | 3.26 | 1.74 |
| | 3mo | 2.10 | 2.10 | 1.82 | 0.52 | 0.05 |
| | 6mo | 0.56 | 1.51 | 1.48 | 0.70 | -0.10 |
| | 12mo | -1.22 | 0.89 | 2.70 | 1.17 | 1.05 |
| h=5 | 1mo | 1.81 | 1.40 | 1.09 | 0.86 | 0.53 |
| | 2mo | 0.91 | 0.72 | 0.53 | 0.41 | 0.27 |
| | 3mo | -0.07 | 0.14 | -0.09 | -0.31 | -0.26 |
| | 6mo | -1.79 | -0.63 | -0.15 | -0.31 | -0.67 |
| | 12mo | -2.83 | -0.58 | 2.39 | 1.03 | -0.08 |
| h=10 | 1mo | -0.56 | -0.79 | -0.84 | -0.66 | -0.51 |
| | 2mo | -0.92 | -0.86 | -0.66 | -0.65 | -0.47 |
| | 3mo | -1.10 | -1.12 | -1.07 | -0.84 | -0.49 |
| | 6mo | -2.45 | -1.86 | -0.95 | -0.61 | -0.82 |
| | 12mo | -3.34 | -2.01 | 0.79 | 0.00 | -0.55 |
| h=15 | 1mo | -0.20 | -0.13 | 0.07 | 0.29 | 0.54 |
| | 2mo | -0.76 | -0.56 | -0.25 | 0.11 | 0.48 |
| | 3mo | -1.41 | -1.16 | -0.66 | -0.29 | 0.05 |
| | 6mo | -2.98 | -2.37 | -1.22 | -0.69 | -0.52 |
| | 12mo | -4.17 | -3.42 | -0.74 | -0.81 | -0.64 |
| h=20 | 1mo | -0.99 | -0.86 | -0.44 | -0.13 | 0.22 |
| | 2mo | -1.54 | -1.29 | -0.71 | -0.27 | 0.20 |
| | 3mo | -2.13 | -1.89 | -1.21 | -0.60 | -0.10 |
| | 6mo | -4.20 | -3.57 | -2.19 | -1.19 | -0.68 |
| | 12mo | -5.60 | -5.02 | -2.26 | -1.51 | -0.89 |
| h=25 | 1mo | -1.60 | -1.40 | -1.00 | -0.64 | -0.27 |
| | 2mo | -2.41 | -2.08 | -1.45 | -0.89 | -0.38 |
| | 3mo | -3.09 | -2.77 | -1.93 | -1.10 | -0.47 |
| | 6mo | -5.42 | -4.65 | -2.98 | -1.76 | -1.11 |
| | 12mo | -7.49 | -6.61 | -3.45 | -2.26 | -1.47 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance is coded by color and font thickness and italics:

| | | | | |
|---|---|---|---|---|
| In favor of indicated model: | 1% | 5% | 10% | 20% |
| In favor of random walk model: | *1%* | *5%* | *10%* | *20%* |

**Figure A.3**: Continued

E: FXVIX1 model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -3.74 | -3.58 | -4.49 | -4.69 | -3.96 |
| | 2mo | 2.56 | 2.94 | 2.24 | 1.08 | 0.75 |
| | 3mo | 0.64 | 0.82 | 0.23 | -0.90 | -1.17 |
| | 6mo | -0.83 | 1.40 | 1.35 | -0.12 | -0.60 |
| | 12mo | -1.52 | 2.63 | 1.61 | 0.85 | -0.52 |
| h=2 | 1mo | -2.33 | -2.61 | -3.40 | -4.17 | -4.01 |
| | 2mo | 1.03 | 1.33 | 0.22 | -0.90 | -1.31 |
| | 3mo | -0.25 | -0.65 | -1.21 | -2.36 | -2.50 |
| | 6mo | -0.70 | -0.55 | -0.71 | -1.85 | -2.16 |
| | 12mo | -1.80 | -0.61 | -1.15 | -1.51 | -2.18 |
| h=5 | 1mo | 2.08 | 1.70 | 0.45 | -1.14 | -1.61 |
| | 2mo | -0.16 | 0.06 | -0.99 | -1.54 | -1.78 |
| | 3mo | -1.72 | -1.26 | -1.46 | -2.27 | -2.24 |
| | 6mo | -3.01 | -2.86 | -2.71 | -2.16 | -2.03 |
| | 12mo | -2.71 | -2.68 | -2.74 | -2.06 | -1.96 |
| h=10 | 1mo | 0.23 | 0.70 | -0.13 | -1.27 | -1.49 |
| | 2mo | -2.12 | -1.44 | -1.53 | -2.12 | -1.82 |
| | 3mo | -3.36 | -2.64 | -2.11 | -2.22 | -1.66 |
| | 6mo | -3.59 | -3.32 | -3.05 | -1.60 | -1.32 |
| | 12mo | -3.72 | -3.94 | -3.09 | -1.80 | -1.35 |
| h=15 | 1mo | -2.33 | -0.57 | -0.87 | -1.80 | -1.61 |
| | 2mo | -4.25 | -2.67 | -2.22 | -2.46 | -2.00 |
| | 3mo | -4.62 | -3.65 | -2.58 | -2.30 | -1.65 |
| | 6mo | -4.54 | -4.01 | -3.44 | -1.54 | -1.14 |
| | 12mo | -5.39 | -5.69 | -3.89 | -2.01 | -1.27 |
| h=20 | 1mo | -4.90 | -3.20 | -2.93 | -3.49 | -2.79 |
| | 2mo | -6.44 | -5.04 | -4.00 | -3.50 | -2.64 |
| | 3mo | -6.11 | -5.16 | -3.69 | -2.84 | -1.95 |
| | 6mo | -6.04 | -5.46 | -4.55 | -1.60 | -0.91 |
| | 12mo | -6.30 | -7.15 | -5.04 | -2.24 | -1.16 |
| h=25 | 1mo | -6.99 | -5.17 | -4.90 | -5.02 | -3.98 |
| | 2mo | -7.90 | -6.93 | -5.57 | -4.60 | -3.38 |
| | 3mo | -7.62 | -6.45 | -4.83 | -3.58 | -2.50 |
| | 6mo | -7.58 | -6.92 | -5.76 | -1.98 | -1.08 |
| | 12mo | -7.82 | -8.55 | -6.39 | -2.90 | -1.63 |
| | | 10 | 25 | atm | 75 | 90 |

F: AVEL1 model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -0.45 | -2.21 | -1.08 | -2.19 | -1.67 |
| | 2mo | 1.27 | 0.95 | 0.13 | -0.33 | -0.63 |
| | 3mo | -0.38 | 0.49 | 0.69 | -1.14 | -1.93 |
| | 6mo | -0.75 | 0.31 | 0.17 | -1.43 | -2.22 |
| | 12mo | -0.84 | 1.18 | 0.83 | -0.9 | -2.09 |
| h=2 | 1mo | -0.48 | -1.71 | -0.78 | -2.31 | -1.64 |
| | 2mo | 2.1 | 1.47 | 0.35 | -0.56 | -0.94 |
| | 3mo | 0.9 | 1.05 | 0.88 | -0.98 | -1.39 |
| | 6mo | 1.08 | 0.93 | 0.58 | -1.7 | -1.99 |
| | 12mo | 0.25 | 0.84 | 0.99 | -2.04 | -3.03 |
| h=5 | 1mo | -1.26 | -0.29 | -0.96 | -0.56 | -0.65 |
| | 2mo | 1.08 | 0.9 | -0.89 | -0.55 | -0.73 |
| | 3mo | -0.18 | 0 | -0.38 | -1.02 | -0.99 |
| | 6mo | -0.4 | -0.29 | -0.25 | -3.2 | -2.23 |
| | 12mo | -0.82 | -0.47 | 1.14 | -3.84 | -3.9 |
| h=10 | 1mo | -1.59 | 0.28 | -1.37 | -0.52 | -0.42 |
| | 2mo | 1.4 | 1.64 | -0.98 | -0.28 | -0.37 |
| | 3mo | -0.17 | 0.25 | -0.55 | -0.71 | -0.52 |
| | 6mo | -0.44 | -0.55 | -0.34 | -4.23 | -2.55 |
| | 12mo | -0.78 | -0.8 | 1.6 | -5.68 | -4.66 |
| h=15 | 1mo | -1.81 | 0.09 | -1.83 | -0.32 | -0.37 |
| | 2mo | 1.33 | 1.71 | -1.34 | -0.29 | -0.34 |
| | 3mo | -0.23 | 0.32 | -0.95 | -0.66 | -0.41 |
| | 6mo | -0.41 | -0.46 | 0.31 | -4.78 | -2.74 |
| | 12mo | -0.71 | -0.49 | 3.19 | -5.57 | -4.81 |
| h=20 | 1mo | -2.06 | -0.14 | -2.01 | -0.16 | -0.25 |
| | 2mo | 1.06 | 1.79 | -1.38 | -0.15 | -0.21 |
| | 3mo | -0.24 | 0.38 | -0.78 | -0.5 | -0.29 |
| | 6mo | -0.32 | -0.3 | 1.04 | -5.14 | -2.91 |
| | 12mo | -0.82 | -0.01 | 4.71 | -5.3 | -5.07 |
| h=25 | 1mo | -2.29 | -0.51 | -2.39 | -0.14 | -0.2 |
| | 2mo | 0.96 | 1.82 | -1.5 | -0.18 | -0.18 |
| | 3mo | -0.17 | 0.42 | -0.54 | -0.41 | -0.22 |
| | 6mo | -0.28 | -0.2 | 1.58 | -5.1 | -3.12 |
| | 12mo | -0.83 | 0.15 | 5.63 | -5 | -5.22 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| 1% | 5% | 10% | 20% |
|---|---|---|---|

**Table A.3**: Aggregate out-of-sample R2 and MDM/MCW tests, and summary of separate MDM/MCW and sign tests: Canadian Dollar 2013-2015

| Type of Model | Forecast horizon $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Share preferred[a] Model | R.Walk | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| ARMA | 0.030 *** | -0.025 | -0.116 | -0.248 | -0.416 | -0.540 | -0.625 | 15.43 | n.a. | 2.29 |
| ARIMA | 0.039 *** | 0.007 *** | -0.009 | -0.009 | -0.004 | -0.003 | -0.006 | 30.29 | n.a. | 11.43 |
| CFC3 | 0.094 *** | 0.029 * | -0.036 ++ | -0.052 + | -0.049 | -0.055 | -0.072 | 12.57 | 19.43 | 12.00 |
| CFC3D | 0.159 *** | 0.070 *** | -0.046 +++ | -0.067 ++ | -0.061 | -0.065 | -0.089 | 19.43 | 22.86 | 17.71 |
| CFL3 | -0.037 | 0.000 | -0.027 ++ | -0.023 | -0.005 | -0.004 | -0.013 | 9.71 | 20.57 | 12.00 |
| CFL10 | 0.096 *** | 0.047 *** | -0.010 | -0.014 | 0.002 | 0.001 | -0.008 | 17.71 | 12.00 | 21.71 |
| GG | 0.040 | 0.009 | -0.037 ++ | -0.057 | -0.060 | -0.076 | -0.104 | 22.29 | 35.43 | 21.14 |
| DI5 | 0.068 ** | 0.021 | -0.011 | -0.020 | -0.030 | -0.043 | -0.061 | 7.43 | 5.14 | 3.43 |
| DI10 | 0.097 *** | 0.053 ** | -0.010 | -0.020 | -0.031 | -0.047 | -0.063 | 10.86 | 2.86 | 12.00 |
| DI5F | 0.069 ** | -0.027 | -0.043 | -0.029 | -0.027 | -0.052 | -0.070 | 4.57 | 14.86 | 4.57 |
| DI10F | 0.064 * | -0.033 | -0.054 | -0.034 | -0.030 | -0.059 | -0.073 | 2.86 | 16.00 | 5.14 |
| DI1FD | 0.160 *** | 0.072 *** | -0.036 ++ | -0.053 + | -0.049 | -0.055 | -0.072 | 20.00 | 21.71 | 13.14 |
| DI5FD | 0.153 *** | 0.048 | -0.023 | -0.019 | -0.016 | -0.047 | -0.070 | 12.00 | 1.71 | 12.57 |
| DI10FD | 0.150 *** | 0.054 * | -0.014 | -0.012 | -0.015 | -0.046 | -0.069 | 10.86 | 1.71 | 12.57 |
| AR2p5 | 0.055 *** | 0.006 ** | -0.046 | -0.065 | -0.060 | -0.064 | -0.079 | 24.57 | n.a. | 9.71 |
| AR2p5D | 0.150 *** | 0.077 *** | -0.023 | -0.040 | -0.042 | -0.051 | -0.070 | 27.43 | n.a. | 10.86 |
| SARpdP | 0.051 | 0.002 | -0.038 | -0.055 | -0.050 | -0.056 | -0.073 | 23.43 | n.a. | 5.14 |
| SAR211 | 0.058 | 0.008 | -0.046 | -0.064 | -0.058 | -0.063 | -0.078 | 24.57 | n.a. | 9.14 |
| FXVIXC | 0.129 *** | 0.060 *** | -0.025 | -0.048 | -0.047 | -0.054 | -0.073 | 22.86 | n.a. | 8.57 |
| AVEC | 0.162 *** | 0.082 *** | -0.021 | -0.037 | -0.039 | -0.049 | -0.069 | 28.57 | n.a. | 12.00 |
| FXVIXL1 | 0.015 | -0.021 | -0.035 | -0.090 | -0.144 | -0.148 | -0.135 | 8.00 | n.a. | 1.14 |
| FXVIXL2 | 0.005 ** | 0.006 ** | -0.015 | -0.048 | -0.094 | -0.099 | -0.095 | 21.71 | n.a. | 6.86 |
| AVEL1 | 0.095 | 0.046 | -0.010 | -0.009 | 0.006 | 0.006 | -0.001 | 34.29 | n.a. | 16.57 |
| AVEL2 | 0.089 | 0.037 | -0.013 | -0.013 | 0.001 | 0.003 | -0.004 | 34.29 | n.a. | 12.57 |

Notes: [a] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable. Significance levels: one-sided tests in favor of indicated model: *** 1%, ** 5%, * 10%; one-sided tests in favor of random walk model: +++ 1%, ++ 5%, + 10%.

**Figure A.4**: Percentage difference in RMSPE relative to random walk, by forecast horizon h, delta and maturity: Canadian dollar, 2013-2015

A: ARIMA model

| h | Maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -0.48 | -1.60 | 1.32 | -0.63 | -1.67 |
| | 2mo | -3.40 | -3.68 | -3.55 | -3.31 | -2.88 |
| | 3mo | -2.15 | -2.58 | -2.60 | -2.68 | -2.15 |
| | 6mo | -3.13 | -3.39 | -3.93 | -3.63 | -3.27 |
| | 12mo | -2.22 | -2.76 | -2.95 | -2.62 | -2.27 |
| h=2 | 1mo | -1.53 | -1.04 | -0.35 | 0.08 | -0.80 |
| | 2mo | **0.03** | 0.31 | 0.34 | 0.48 | 0.53 |
| | 3mo | 0.17 | **0.11** | **-0.02** | **-0.10** | 0.19 |
| | 6mo | -0.63 | -0.72 | -0.87 | -0.97 | -0.92 |
| | 12mo | -0.61 | -0.83 | -0.77 | -0.80 | -0.63 |
| h=5 | 1mo | **-0.64** | -0.50 | 0.07 | 0.55 | 0.65 |
| | 2mo | 0.68 | 0.79 | 0.96 | 0.99 | 0.97 |
| | 3mo | 0.72 | 0.91 | 0.94 | 0.93 | 0.74 |
| | 6mo | 0.35 | 0.51 | 0.60 | 0.65 | 0.37 |
| | 12mo | -0.02 | 0.20 | 0.31 | 0.27 | -0.01 |
| h=10 | 1mo | -0.37 | -0.83 | -0.58 | -0.25 | 0.15 |
| | 2mo | 0.70 | 0.80 | 0.90 | 1.07 | 1.16 |
| | 3mo | 0.73 | 0.94 | 1.02 | 1.10 | 0.82 |
| | 6mo | 0.66 | 0.86 | 1.06 | 1.17 | 1.03 |
| | 12mo | 0.25 | 0.49 | 0.67 | 0.77 | 0.59 |
| h=15 | 1mo | -0.06 | -0.24 | **-0.54** | 0.00 | -0.07 |
| | 2mo | 0.31 | 0.35 | 0.37 | 0.53 | 0.52 |
| | 3mo | 0.29 | 0.42 | 0.50 | 0.56 | 0.16 |
| | 6mo | 0.33 | 0.48 | 0.66 | 0.76 | 0.71 |
| | 12mo | 0.13 | 0.30 | 0.48 | 0.57 | 0.47 |
| h=20 | 1mo | 0.35 | -0.04 | -0.40 | 0.01 | -0.39 |
| | 2mo | 0.26 | 0.27 | 0.32 | 0.41 | 0.43 |
| | 3mo | 0.23 | 0.34 | 0.40 | 0.48 | 0.06 |
| | 6mo | 0.17 | 0.27 | 0.38 | 0.49 | 0.48 |
| | 12mo | 0.03 | 0.13 | 0.24 | 0.36 | 0.33 |
| h=25 | 1mo | 0.36 | 0.02 | -0.05 | 0.17 | -0.07 |
| | 2mo | 0.35 | 0.40 | 0.43 | 0.53 | 0.54 |
| | 3mo | 0.27 | 0.40 | 0.48 | 0.58 | 0.45 |
| | 6mo | 0.20 | 0.36 | 0.50 | 0.65 | 0.62 |
| | 12mo | 0.10 | 0.26 | 0.40 | 0.54 | 0.47 |

B: AVEC model

| h | Maturity | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -13.37 | -13.21 | -13.30 | -12.70 | -12.06 |
| | 2mo | -8.46 | -8.41 | -8.29 | -7.35 | -6.41 |
| | 3mo | -3.37 | -3.78 | -3.72 | -3.38 | -2.98 |
| | 6mo | -3.57 | -3.81 | -4.07 | -3.61 | -3.33 |
| | 12mo | -2.72 | -2.76 | -2.71 | -2.29 | -1.97 |
| h=2 | 1mo | -7.67 | -7.37 | -7.11 | -6.77 | -6.15 |
| | 2mo | -5.43 | -5.17 | -4.89 | -4.11 | -3.35 |
| | 3mo | -0.75 | -0.80 | -0.73 | -0.40 | -0.06 |
| | 6mo | -0.95 | -1.01 | -0.97 | -0.74 | -0.50 |
| | 12mo | -0.48 | -0.24 | 0.04 | 0.24 | **0.34** |
| h=5 | 1mo | 0.76 | 0.80 | 0.76 | 0.77 | 0.81 |
| | 2mo | 1.27 | 1.28 | 1.23 | 1.22 | 1.27 |
| | 3mo | 1.49 | 1.47 | 1.45 | 1.46 | 1.53 |
| | 6mo | 1.17 | 1.08 | 1.02 | 1.11 | 1.16 |
| | 12mo | 0.92 | 0.96 | 0.92 | 0.93 | 0.81 |
| h=10 | 1mo | 1.43 | 1.47 | 1.48 | 1.55 | 1.65 |
| | 2mo | 1.80 | 1.83 | 1.80 | 1.91 | 1.97 |
| | 3mo | 2.03 | 2.06 | 2.07 | 2.20 | 2.36 |
| | 6mo | 2.16 | 2.11 | 2.12 | 2.29 | 2.46 |
| | 12mo | 1.85 | 1.92 | 1.91 | 2.05 | 2.01 |
| h=15 | 1mo | 1.48 | 1.48 | 1.44 | 1.54 | 1.67 |
| | 2mo | 1.99 | 1.88 | 1.77 | 1.84 | 1.90 |
| | 3mo | 2.22 | 2.11 | 2.06 | 2.15 | 2.34 |
| | 6mo | 2.53 | 2.33 | 2.30 | 2.44 | 2.67 |
| | 12mo | 2.39 | 2.29 | 2.28 | 2.43 | 2.53 |
| h=20 | 1mo | 2.07 | 2.02 | 1.96 | 2.05 | 2.22 |
| | 2mo | 2.55 | 2.36 | 2.25 | 2.29 | 2.40 |
| | 3mo | 2.72 | 2.54 | 2.45 | 2.55 | 2.74 |
| | 6mo | 3.04 | 2.71 | 2.62 | 2.75 | 3.05 |
| | 12mo | 3.02 | 2.71 | 2.69 | 2.88 | 3.18 |
| h=25 | 1mo | 2.99 | 2.92 | 2.83 | 2.91 | 3.08 |
| | 2mo | 3.58 | 3.36 | 3.20 | 3.25 | 3.35 |
| | 3mo | 3.75 | 3.55 | 3.43 | 3.51 | 3.72 |
| | 6mo | 4.02 | 3.64 | 3.54 | 3.69 | 4.06 |
| | 12mo | 3.88 | 3.59 | 3.59 | 3.84 | 4.22 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| 1% | 5% | 10% | 20% |
|---|---|---|---|

**Figure A.4**: Continued

C: CFL10 model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -8.24 | -8.14 | -7.33 | -7.55 | -7.42 |
| | 2mo | -5.02 | -4.53 | -4.56 | -4.72 | -4.72 |
| | 3mo | -0.79 | 0.10 | -0.65 | -2.20 | -2.42 |
| | 6mo | -1.00 | -1.62 | -1.56 | -2.11 | -3.15 |
| | 12mo | -0.22 | -1.17 | -1.88 | -3.14 | -4.57 |
| h=2 | 1mo | -6.43 | -6.34 | -5.44 | -5.44 | -4.87 |
| | 2mo | -2.85 | -2.29 | -1.92 | -2.06 | -1.88 |
| | 3mo | 0.83 | 2.02 | 2.18 | 1.34 | 1.41 |
| | 6mo | 0.49 | 0.81 | 1.29 | 1.31 | 1.09 |
| | 12mo | 1.73 | 1.54 | 1.53 | 0.88 | 0.07 |
| h=5 | 1mo | -1.91 | -2.38 | -2.18 | -2.12 | -2.20 |
| | 2mo | 1.50 | 1.83 | 2.14 | 2.21 | 2.38 |
| | 3mo | 3.12 | 3.68 | 3.56 | 2.92 | 2.71 |
| | 6mo | 1.59 | 1.20 | 1.02 | 1.13 | 0.81 |
| | 12mo | 2.45 | 1.23 | 0.55 | 0.13 | -0.60 |
| h=10 | 1mo | -2.28 | -2.58 | -2.28 | -2.00 | -1.93 |
| | 2mo | 1.32 | 1.91 | 2.28 | 2.48 | 2.93 |
| | 3mo | 3.31 | 3.73 | 3.38 | 2.64 | 2.58 |
| | 6mo | 2.86 | 1.97 | 1.57 | 1.53 | 1.02 |
| | 12mo | 3.43 | 1.86 | 0.87 | 0.37 | -0.57 |
| h=15 | 1mo | -4.14 | -4.37 | -3.91 | -3.17 | -2.60 |
| | 2mo | 0.20 | 0.78 | 1.18 | 1.79 | 2.57 |
| | 3mo | 2.38 | 3.00 | 2.69 | 2.15 | 2.34 |
| | 6mo | 2.99 | 2.05 | 1.58 | 1.29 | 0.94 |
| | 12mo | 3.85 | 2.01 | 0.78 | 0.06 | -1.08 |
| h=20 | 1mo | -4.79 | -4.86 | -4.12 | -2.70 | -1.53 |
| | 2mo | -0.54 | -0.03 | 0.69 | 1.94 | 3.21 |
| | 3mo | 1.64 | 2.36 | 2.38 | 2.72 | 3.54 |
| | 6mo | 3.21 | 1.86 | 1.62 | 1.86 | 2.07 |
| | 12mo | 3.92 | 1.68 | 0.75 | 0.48 | -0.19 |
| h=25 | 1mo | -4.00 | -4.12 | -3.17 | -1.23 | 0.38 |
| | 2mo | -0.16 | -0.03 | 0.50 | 1.84 | 3.07 |
| | 3mo | 2.48 | 2.93 | 2.83 | 3.37 | 4.13 |
| | 6mo | 3.45 | 1.87 | 1.33 | 1.73 | 1.79 |
| | 12mo | 3.46 | 1.19 | 0.06 | 0.14 | -0.60 |
| | | 10 | 25 | atm | 75 | 90 |

D: DI10FD model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -13.80 | -13.50 | -13.39 | -12.51 | -11.82 |
| | 2mo | -7.85 | -7.62 | -7.05 | -5.97 | -4.83 |
| | 3mo | -2.24 | -2.39 | -2.21 | -1.68 | -1.23 |
| | 6mo | -3.19 | -3.18 | -3.14 | -2.56 | -2.22 |
| | 12mo | -2.13 | -1.89 | -1.64 | -1.46 | -1.57 |
| h=2 | 1mo | -7.74 | -7.34 | -6.85 | -6.17 | -5.25 |
| | 2mo | -4.73 | -3.97 | -3.34 | -2.11 | -0.47 |
| | 3mo | 0.80 | 1.11 | 1.71 | 3.14 | 4.21 |
| | 6mo | 0.37 | 0.79 | 0.93 | 2.01 | 3.21 |
| | 12mo | 1.78 | 2.22 | 2.73 | 3.52 | 4.01 |
| h=5 | 1mo | -0.02 | -0.04 | -0.02 | 0.22 | 0.58 |
| | 2mo | 0.38 | 0.59 | 0.54 | 0.87 | 1.64 |
| | 3mo | 0.90 | 0.87 | 0.99 | 1.55 | 2.22 |
| | 6mo | 0.84 | 0.87 | 0.75 | 1.24 | 2.13 |
| | 12mo | 1.08 | 0.91 | 0.89 | 1.25 | 2.22 |
| h=10 | 1mo | -0.64 | -0.56 | -0.55 | -0.33 | 0.03 |
| | 2mo | 0.57 | 0.77 | 0.75 | 0.93 | 1.34 |
| | 3mo | 0.71 | 0.73 | 0.72 | 1.06 | 1.39 |
| | 6mo | 1.37 | 1.36 | 1.25 | 1.49 | 1.88 |
| | 12mo | 1.80 | 1.56 | 1.46 | 1.69 | 2.13 |
| h=15 | 1mo | -0.01 | -0.03 | -0.11 | 0.03 | 0.33 |
| | 2mo | 1.05 | 0.93 | 0.77 | 0.72 | 0.82 |
| | 3mo | 1.33 | 1.07 | 0.86 | 0.91 | 1.01 |
| | 6mo | 1.62 | 1.35 | 1.13 | 1.14 | 1.22 |
| | 12mo | 1.99 | 1.63 | 1.40 | 1.42 | 1.69 |
| h=20 | 1mo | 1.70 | 1.69 | 1.64 | 1.81 | 2.09 |
| | 2mo | 2.33 | 2.20 | 2.05 | 2.08 | 2.29 |
| | 3mo | 2.90 | 2.66 | 2.54 | 2.64 | 2.70 |
| | 6mo | 3.08 | 2.70 | 2.44 | 2.44 | 2.60 |
| | 12mo | 3.36 | 2.95 | 2.76 | 2.72 | 2.98 |
| h=25 | 1mo | 2.55 | 2.45 | 2.32 | 2.45 | 2.74 |
| | 2mo | 3.58 | 3.36 | 3.07 | 3.15 | 3.46 |
| | 3mo | 4.35 | 4.06 | 3.86 | 3.97 | 4.12 |
| | 6mo | 4.58 | 4.13 | 3.78 | 3.82 | 4.10 |
| | 12mo | 4.70 | 4.24 | 4.07 | 4.09 | 4.41 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance is coded by color and font thickness and italics:

| | | | | |
|---|---|---|---|---|
| In favor of indicated model: | 1% | 5% | 10% | 20% |
| In favor of random walk model: | *1%* | *5%* | *10%* | *20%* |

**Figure A.4**: Continued

E: FXVIX1 model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -1.41 | -0.03 | -0.19 | -0.23 | 0.63 |
| | 2mo | -2.40 | -0.13 | -0.83 | -0.17 | -0.05 |
| | 3mo | -0.20 | 0.14 | 0.28 | 0.34 | 0.44 |
| | 6mo | -1.17 | -0.87 | 0.16 | 0.38 | 0.69 |
| | 12mo | 0.49 | 0.25 | 0.71 | 0.82 | 0.17 |
| h=2 | 1mo | -2.35 | -0.93 | -1.23 | -0.96 | -0.47 |
| | 2mo | -1.46 | -0.05 | -0.30 | 0.05 | 0.23 |
| | 3mo | -0.23 | 0.19 | 0.62 | 0.75 | 1.06 |
| | 6mo | -0.13 | 0.14 | 1.08 | 1.43 | 1.73 |
| | 12mo | 1.56 | 1.07 | 1.64 | 1.71 | 1.39 |
| h=5 | 1mo | 0.52 | -0.68 | -1.39 | -0.95 | -0.90 |
| | 2mo | -0.20 | 0.21 | 1.22 | 1.79 | 2.29 |
| | 3mo | 0.07 | 1.02 | 1.40 | 2.04 | 2.30 |
| | 6mo | 0.16 | 1.42 | 3.30 | 3.53 | 3.10 |
| | 12mo | 3.31 | 2.72 | 2.30 | 2.18 | 1.74 |
| h=10 | 1mo | 2.59 | 0.34 | -0.12 | 0.46 | 1.39 |
| | 2mo | -0.83 | 0.79 | 2.43 | 3.81 | 5.26 |
| | 3mo | -0.30 | 3.34 | 2.79 | 4.04 | 4.82 |
| | 6mo | 0.56 | 1.98 | 5.75 | 6.34 | 6.08 |
| | 12mo | 6.22 | 4.80 | 4.12 | 4.26 | 3.97 |
| h=15 | 1mo | 3.93 | 2.22 | 2.10 | 3.12 | 4.71 |
| | 2mo | -0.82 | 0.37 | 4.16 | 6.63 | 9.21 |
| | 3mo | -0.32 | 5.73 | 5.47 | 8.05 | 9.80 |
| | 6mo | 1.01 | 4.26 | 9.41 | 10.91 | 11.14 |
| | 12mo | 6.98 | 5.50 | 6.54 | 7.37 | 7.22 |
| h=20 | 1mo | 3.09 | 3.58 | 2.59 | 3.99 | 5.62 |
| | 2mo | -0.89 | -0.48 | 3.98 | 6.27 | 9.06 |
| | 3mo | -0.33 | 5.91 | 5.87 | 8.47 | 10.34 |
| | 6mo | 0.86 | 5.91 | 9.99 | 11.69 | 12.21 |
| | 12mo | 5.94 | 5.22 | 6.87 | 7.79 | 7.78 |
| h=25 | 1mo | 3.42 | 4.09 | 3.02 | 4.48 | 6.46 |
| | 2mo | -1.08 | -1.71 | 2.76 | 4.68 | 7.51 |
| | 3mo | -0.66 | 5.91 | 5.45 | 7.97 | 10.07 |
| | 6mo | 0.59 | 6.18 | 9.31 | 10.94 | 11.75 |
| | 12mo | 5.13 | 6.13 | 6.04 | 7.22 | 7.32 |
| | | 10 | 25 | atm | 75 | 90 |

F: AVEL1 model

| h | | 10 | 25 | atm | 75 | 90 |
|---|---|---|---|---|---|---|
| h=1 | 1mo | -7.21 | -7.61 | -8.10 | -6.88 | -6.33 |
| | 2mo | -4.98 | -5.30 | -5.67 | -5.49 | -5.26 |
| | 3mo | -0.53 | -1.43 | -0.84 | -1.90 | -1.24 |
| | 6mo | -1.05 | -1.11 | -2.18 | -1.91 | -2.46 |
| | 12mo | -0.55 | -1.69 | -2.58 | -2.99 | -2.45 |
| h=2 | 1mo | -5.05 | -5.48 | -5.91 | -4.66 | -3.97 |
| | 2mo | -2.13 | -2.39 | -2.72 | -2.60 | -2.34 |
| | 3mo | 1.60 | 0.77 | 1.49 | 0.60 | 1.54 |
| | 6mo | 1.11 | 1.21 | 0.34 | 0.23 | 0.20 |
| | 12mo | 1.52 | 1.04 | 0.14 | -0.41 | -0.17 |
| h=5 | 1mo | -0.43 | -1.24 | -2.05 | -1.35 | -0.30 |
| | 2mo | 1.56 | 1.35 | 1.33 | 1.28 | 0.69 |
| | 3mo | 2.28 | 2.04 | 2.68 | 1.49 | 0.97 |
| | 6mo | 1.78 | 2.33 | 1.11 | 0.27 | 0.06 |
| | 12mo | 1.83 | 1.98 | 0.64 | -0.04 | -0.23 |
| h=10 | 1mo | -1.56 | -2.47 | -2.98 | -1.26 | 0.33 |
| | 2mo | 1.30 | 1.21 | 1.15 | 1.78 | 1.21 |
| | 3mo | 1.75 | 1.77 | 2.02 | 1.85 | 1.35 |
| | 6mo | 1.33 | 3.01 | 1.93 | 1.18 | 0.90 |
| | 12mo | 1.72 | 2.92 | 1.36 | 0.56 | 0.50 |
| h=15 | 1mo | -3.07 | -4.09 | -4.48 | -1.92 | -0.28 |
| | 2mo | 0.64 | 0.39 | 0.11 | 1.43 | 0.59 |
| | 3mo | 0.95 | 0.99 | 1.27 | 1.47 | 0.67 |
| | 6mo | 0.65 | 2.87 | 2.12 | 0.60 | 0.57 |
| | 12mo | 1.28 | 3.26 | 1.56 | 0.28 | 0.23 |
| h=20 | 1mo | -3.29 | -4.34 | -4.40 | -1.40 | 0.59 |
| | 2mo | 0.47 | 0.26 | -0.11 | 1.42 | 0.29 |
| | 3mo | 0.64 | 0.64 | 1.31 | 1.34 | 0.30 |
| | 6mo | 0.63 | 2.90 | 2.20 | 0.26 | 0.23 |
| | 12mo | 1.40 | 3.05 | 2.37 | 0.11 | 0.04 |
| h=25 | 1mo | -3.43 | -4.29 | -3.75 | -0.40 | 1.79 |
| | 2mo | 0.64 | 0.53 | -0.35 | 1.63 | 0.42 |
| | 3mo | 0.87 | 0.88 | 1.89 | 2.33 | 0.43 |
| | 6mo | 0.77 | 2.83 | 2.22 | 0.18 | 0.28 |
| | 12mo | 0.98 | 2.49 | 3.01 | 0.15 | 0.07 |
| | | 10 | 25 | atm | 75 | 90 |

Statistical significance in favor of the indicated model is coded by color and font thickness:

| 1% | 5% | 10% | 20% |
|---|---|---|---|

## A.2 The fifth group of models: VAR

This appendix briefly describes VAR models that could be used for forecasting option volatility. In the literature, Guo et al (2018) list them among their model alternatives. These models constitute the fifth group, in addition to the four groups described in Section 3. As shown below, VAR models may be specified in levels and in first difference.

V.1 VARC: multivariate VAR(1) in first difference

$$\Delta y_t = \alpha + A\Delta y_{t-1} + \epsilon_t \tag{A2.1}$$

When forecasting, this model is rolled forward.

V.2 VARL: multivariate VAR(1) in levels

$$y_t = \alpha + Ay_{t-1} + \epsilon_t \tag{A2.2}$$

When forecasting, this model is rolled forward.

V.3 VAR*C: multivariate VAR(1*) in first difference

This forecasting strategy consists of a collection of models, one for each forecasting window $h$.

$$\Delta_h y_t = \alpha_h + A_h \Delta_h y_{t-h} + \epsilon_{th} \tag{A2.3}$$

The forecasting equation follows immediately:

$$\widehat{\Delta_h \mathrm{y}}_{T+h} = \hat{\alpha}_h + \hat{A}_h \Delta_h y_T$$

and thus

$$\hat{y}_{T+h} = y_T + \hat{A}_h \widehat{\Delta_h \mathrm{y}}_{T+h}$$

V.4 VAR*Cm: multivariate VAR(1*) on first-differenced options with the same maturity

The VARC and VARL models are highly parameter-intensive, despite the fact that the contain only one lag: in equations (A2.1)-(A2.1), the matrix $A$ contains $25^2 = 625$ parameters. The VAR*C and VAR*L models are even worse in this regard, as each forecast window $h$ requires an estimate of a different matrix $A_h$ with 625 parameters to forecast a $25 \times 1$ vector $y_{T+h}$. This may work in other environments, but the movements of the $25$ options are highly correlated and therefore may generate high degrees of multicollinearity in the estimated equations, imprecisely estimated parameters, and noisy forecasts. Guo et al (2018) propose to forecast with smaller VAR* models aimed at options with the same maturity. Let $y_{mt} = \left(y_{10,mt}, y_{25,mt}, y_{50,mt}, y_{75,mt}, y_{90,mt}\right)$ be the $5 \times 1$ vector that stacks such options for maturity $m$. The VAR*Cm model is a VAR(1*) model of $\Delta y_{mt}$:

$$\Delta_h y_{mt} = \alpha_{mh} + A_{mh} \Delta_h y_{m,t-h} + \epsilon_{mt} \tag{A2.4}$$

This is a special case of the VAR*C model in equation (A2.3) with $A_h = diag(A_{mh})$, a block-diagonal matrix, where off-diagonal blocks are restricted to 0. This reduces the number of parameters in $A_h$ from 625 to 125.

V.5 VAR*Cd: multivariate VAR(1*) on first-differenced options with the same delta

An alternative to VAR*Cm is to focus on options with the same degree of moneyness. Thus, let $y_{dt} = \left(y_{d,1,t}, y_{d,2,t}, y_{d,3,t}, y_{d,6,t}, y_{d,12,t}\right)$ stack such options.

$$\Delta_h y_{dt} = \alpha_{dh} + A_{dh} \Delta_h y_{d,t-h} + \epsilon_{dth} \tag{A2.5}$$

If $y_{dmt}$ is stacked into $y_t$ with $d$ slow and $m$ fast, $A_h$ may once again be written as a block-diagonal matrix $diag(A_{dh})$. Since $y_{dmt}$ is arguably related to both $y_{d'mt}$ for $d \neq d'$ and $y_{dm't}$ for $m \neq m'$, it is an empirical question whether the restriction imposed by the VAR*Cm model yields better forecasts than VAR*Cd—and whether either one is better than VAR*C.

V.6 VAR*L: multivariate VAR(1*) in levels

$$y_t = \alpha_h + A_h y_{t-h} + \epsilon_{th} \tag{A2.6}$$

Once again, the forecasting equation follows immediately:

$$\hat{y}_{T+h} = \hat{\alpha}_h + \hat{A}_h y_T$$

V.7 VAR*Lm: multivariate VAR(1*) on options (in levels) with the same maturity

$$y_{mt} = \alpha_{mh} + A_{mh} y_{m,t-h} + \epsilon_{mth} \tag{A2.7}$$

V.8 VAR*Ld: multivariate VAR(1*) on options (in levels) with the same delta

$$y_{dt} = \alpha_{dh} + A_{dh} y_{d,t-h} + \epsilon_{dth} \tag{A2.8}$$

Table A.4 reports the aggregate tests of these VAR models for Euro implied volatility for both 2010-2012 and 2013-2015. The VARC model provides a case study of anomalies in these tests: $R^2_{OS}$ is negative in all cases, indicating that VARC forecast errors are generally larger than random walk forecast errors. Yet the MCW test overwhelmingly rejects the hypothesis that the random walk and VARC models are equivalent, in favor of the alternative hypothesis that the VARC at the true parameters is a better forecasting model. This is because the VARC model is very parameter-intensive and thus contains numerous imprecisely estimated parameters that adversely affect the small-sample quality of the forecasts. In other words, in very large samples, VAR1C should outperform the random walk, but at least in the present case the small sample performance is weak. The same comment applies to the VARL model, which generally performs worse yet than VARC. The VAR*C and VAR*L models are yet more parameter-intensive and perform yet worse, judged by the $R^2_{OS}$. The restrictions imposed by the VAR*Cm and VAR*Cd models are somewhat useful—it appears to be better to group the option volatility time series by delta than by maturity—but these models may outperform the random walk model only at the one- or two-day forecast horizon. Similarly, the $R^2_{OS}$ values of VAR*Lm and VAR*Ld model is better than VARL but still negative at all forecast horizons in 2010-2012.

**Table A.4**: Euro implied volatility forecasts with VAR models

| Type of Model | Forecast horizon $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Share preferred[a] Model | R.Walk | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| A: Euro, 2010-2012 | | | | | | | | | | |
| VARC | -0.108 *** | -0.058 *** | -0.016 *** | -0.023 *** | -0.024 *** | -0.030 *** | -0.035 *** | 100.00 | n.a. | 17.14 |
| VAR*C | -0.109 | -0.119 ** | -0.412 | -0.266 | -0.334 | -0.687 | -1.092 | 19.43 | n.a. | 0.00 |
| VAR*Cm | -0.012 ** | -0.036 | -0.076 | -0.093 | -0.002 | -0.128 | -0.146 | 16.00 | n.a. | 8.57 |
| VAR*Cd | 0.005 ** | 0.000 ** | -0.106 | -0.089 | -0.115 | -0.229 | -0.357 | 17.71 | n.a. | 1.71 |
| VARL | -0.099 *** | -0.077 *** | -0.090 *** | -0.040 *** | -0.056 *** | -0.170 *** | -0.252 *** | 100.00 | n.a. | 0.57 |
| VAR*L | -0.109 | -0.119 ** | -0.412 | -0.266 | -0.334 | -0.687 | -1.092 | 19.43 | n.a. | 0.00 |
| VAR*Lm | -0.019 * | -0.022 ** | -0.096 | -0.133 | -0.231 | -0.389 | -0.428 | 14.29 | n.a. | 0.00 |
| VAR*Ld | -0.017 ** | -0.032 | -0.087 | -0.182 | -0.398 | -0.531 | -0.723 | 6.29 | n.a. | 0.00 |

| Type of Model | $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Share preferred[a] Model | Share preferred[a] R.Walk | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| B: Euro, 2013-2015 | | | | | | | | | | |
| VARC | -0.067 *** | -0.021 *** | -0.028 *** | -0.031 *** | -0.035 *** | -0.044 *** | -0.054 *** | 100.00 | n.a. | 0.00 |
| VAR*C | -0.069 | -0.058 * | -0.186 | -0.196 | -0.086 ** | -0.171 * | -0.124 ** | 28.57 | n.a. | 8.57 |
| VAR*Cm | -0.011 | -0.010 | -0.046 | -0.053 | -0.093 | -0.061 *** | -0.051 ** | 20.57 | n.a. | 4.57 |
| VAR*Cd | 0.001 ** | 0.020 *** | -0.027 | -0.018 | -0.042 | -0.010 | -0.002 | 21.71 | n.a. | 9.14 |
| VARL | -0.098 *** | -0.105 *** | -0.165 *** | -0.142 *** | -0.143 *** | -0.214 *** | -0.346 *** | 100.00 | n.a. | 0.00 |
| VAR*L | -0.097 | -0.118 * | -0.210 * | -0.349 * | -0.409 ** | -0.694 ** | -0.701 * | 34.29 | n.a. | 0.00 |
| VAR*Lm | -0.027 | -0.045 | -0.110 | -0.220 | -0.288 | -0.379 | -0.464 | 0.00 | n.a. | 0.00 |
| VAR*Ld | 0.006 *** | 0.026 *** | 0.004 * | 0.016 * | 0.006 | 0.024 * | -0.021 | 25.71 | n.a. | 12.57 |

Notes: [a] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable. Significance levels: *** 1%, ** 5%, * 10% (one-sided tests)

**Table A.5**: Canadian Dollar implied volatility forecasts with VAR models

| Type of Model | Forecast horizon | | | | | | | Share preferred[a] | | Sign test |
|---|---|---|---|---|---|---|---|---|---|---|
| | $h=1$ | $h=2$ | $h=5$ | $h=10$ | $h=15$ | $h=20$ | $h=25$ | Model | R.Walk | |
| A: Canadian Dollar, 2010-2012 | | | | | | | | | | |
| VARC | -0.179 *** | -0.036 *** | -0.028 *** | 0.002 *** | 0.012 *** | 0.021 *** | 0.026 *** | 100.00 | n.a. | 52.57 |
| VAR*C | -0.182 | -0.187 | -0.529 | -0.403 | -0.640 | -0.839 | -0.443 | 3.43 | n.a. | 0.00 |
| VAR*Cm | -0.008 ** | -0.029 | -0.140 | -0.166 | -0.232 | -0.274 | -0.105 | 12.57 | n.a. | 1.71 |
| VAR*Cd | -0.008 ** | 0.006 ** | -0.082 | -0.057 | -0.130 | -0.145 | -0.069 | 19.43 | n.a. | 14.29 |
| VARL | -0.133 *** | -0.155 *** | -0.286 *** | -0.337 *** | -0.331 *** | -0.312 *** | -0.305 *** | 100.00 | n.a. | 0.00 |
| VAR*L | -0.132 | -0.307 | -0.743 | -1.032 | -1.539 | -1.542 | -1.340 | 2.29 | n.a. | 0.00 |
| VAR*Lm | -0.016 | -0.055 | -0.133 | -0.223 | -0.307 | -0.388 | -0.460 | 2.29 | n.a. | 1.14 |
| VAR*Ld | -0.013 | -0.042 | -0.134 | -0.254 | -0.440 | -0.599 | -0.684 | 2.86 | n.a. | 0.57 |
| B: Canadian Dollar, 2013-2015 | | | | | | | | | | |
| VARC | 0.054 *** | 0.016 *** | -0.038 *** | -0.051 *** | -0.054 *** | -0.055 *** | -0.067 ** | 100.00 | n.a. | 7.43 |
| VAR*C | 0.057 *** | -0.020 *** | -0.025 * | -0.287 | -0.573 | -0.792 | -0.702 | 24.00 | n.a. | 4.57 |
| VAR*Cm | 0.029 *** | -0.017 | -0.012 | -0.051 | -0.068 | 0.011 | -0.068 | 16.00 | n.a. | 1.14 |
| VAR*Cd | 0.088 *** | -0.003 *** | 0.030 ** | -0.049 | -0.161 | -0.126 | -0.050 | 31.43 | n.a. | 10.29 |
| VARL | -0.020 *** | -0.026 *** | -0.124 *** | -0.243 *** | -0.355 *** | -0.442 ** | -0.513 ** | 100.00 | n.a. | 0.00 |
| VAR*L | -0.018 *** | -0.044 ** | -0.225 | -0.570 | -0.747 | -1.030 | -1.093 | 9.71 | n.a. | 0.00 |
| VAR*Lm | -0.024 | -0.033 | -0.132 | -0.270 | -0.463 | -0.594 | -0.680 | 0.00 | n.a. | 0.00 |
| VAR*Ld | 0.019 *** | 0.002 *** | -0.102 | -0.247 | -0.365 | -0.418 | -0.434 | 10.86 | n.a. | 1.14 |

Notes: [a] Share of 175 forecasted series (25 option volatilities with 7 forecast horizons) for which the specified model (left column) or the random walk model (right column) is preferred in the MDM/MCW test at a 10% significance level. "n.a." indicates that the specified model nests the random walk model, in which case the notion that the random walk model is preferred is not applicable. Significance levels: *** 1%, ** 5%, * 10% (one-sided tests)

### A.3 Detailed results on the rate of return to trading implied volatility

**Table A.6**: Rate of return to trading implied volatility: Euro, 2010-2012

A: No Transaction Cost

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.000 | 5.48$^{***}$ | 4.43$^{***}$ | 4.31$^{***}$ | 6.07$^{***}$ | 3.23$^{**}$ | 5.21$^{***}$ | 4.21$^{***}$ | 5.21$^{***}$ |
| 2 | 0.000 | 5.17$^{***}$ | 3.59$^{***}$ | 4.75$^{***}$ | 4.64$^{***}$ | 3.00$^{***}$ | 3.46$^{***}$ | 3.09$^{***}$ | 4.82$^{***}$ |
| 5 | 0.000 | 0.94 | 0.50 | 0.11 | 1.46$^{*}$ | 3.01$^{***}$ | 1.63$^{*}$ | 3.42$^{***}$ | 3.26$^{***}$ |
| 10 | 0.000 | -0.18 | 0.28 | 0.03 | 0.09 | 2.94$^{***}$ | 1.95$^{**}$ | 3.92$^{***}$ | 3.19$^{***}$ |
| 15 | 0.000 | -0.54 | 0.61 | -0.14 | 0.05 | 2.27$^{**}$ | 1.85$^{*}$ | 3.71$^{***}$ | 3.39$^{***}$ |
| 20 | 0.000 | -0.74 | 0.52 | -0.02 | 0.18 | 1.89 | 1.93$^{*}$ | 3.49$^{***}$ | 3.40$^{***}$ |
| 25 | 0.000 | -0.66 | 0.60 | 0.65 | 0.14 | 1.69 | 1.84$^{*}$ | 3.22$^{**}$ | 3.31$^{***}$ |

B: 2 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -5.45 | 1.96$^{*}_{+++}$ | 1.73$_{+++}$ | 1.60$_{+++}$ | 3.64$^{**}_{+++}$ | 0.24$_{+++}$ | 1.77$_{+++}$ | 1.31$_{+++}$ | 2.29$^{*}_{+++}$ |
| 2 | -5.45 | 1.52$_{+++}$ | 0.92$_{+++}$ | 2.09$^{*}_{+++}$ | 1.95$^{*}_{+++}$ | 0.01$_{+++}$ | 0.55$_{+++}$ | 0.18$_{+++}$ | 1.58$_{+++}$ |
| 5 | -5.45 | -2.77$_{+++}$ | -2.11$_{+++}$ | -3.24$_{+++}$ | -1.33$_{+++}$ | 0.48$_{+++}$ | -1.21$_{+++}$ | -0.02$_{+++}$ | -0.24$_{+++}$ |
| 10 | -5.45 | -3.99$_{+++}$ | -2.56$_{+++}$ | -3.40$_{+++}$ | -3.14$_{++}$ | 0.40$_{+++}$ | -0.78$_{+++}$ | 0.30$_{+++}$ | -0.52$_{+++}$ |
| 15 | -5.45 | -4.41$_{++}$ | -2.63$_{+++}$ | -3.62$_{++}$ | -3.34$_{+}$ | -0.28$_{+++}$ | -0.94$_{+++}$ | -0.04$_{+++}$ | -0.39$_{+++}$ |
| 20 | -5.45 | -4.59 | -2.94$_{+++}$ | -3.58$_{++}$ | -3.28$_{++}$ | -0.59$_{+++}$ | -0.78$_{+++}$ | -0.29$_{+++}$ | -0.52$_{+++}$ |
| 25 | -5.45 | -4.43$_{+}$ | -2.87$_{+++}$ | -3.16$_{+++}$ | -3.33$_{++}$ | -0.79$_{+++}$ | -0.83$_{+++}$ | -0.60$_{+++}$ | -0.79$_{+++}$ |

C: 5 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -13.63 | -3.30$_{+++}$ | -2.31$_{+++}$ | -2.47$_{+++}$ | 0.00$_{+++}$ | -4.24$_{+++}$ | -3.38$_{+++}$ | -3.04$_{+++}$ | -2.10$_{+++}$ |
| 2 | -13.63 | -3.94$_{+++}$ | -3.08$_{+++}$ | -1.89$_{+++}$ | -2.09$_{+++}$ | -4.49$_{+++}$ | -3.81$_{+++}$ | -4.18$_{+++}$ | -3.28$_{+++}$ |
| 5 | -13.63 | -8.34$_{+++}$ | -6.02$_{+++}$ | -8.25$_{+++}$ | -5.51$_{+++}$ | -3.33$_{+++}$ | -5.47$_{+++}$ | -5.19$_{+++}$ | -5.49$_{+++}$ |
| 10 | -13.63 | -9.71$_{+++}$ | -6.82$_{+++}$ | -8.55$_{+++}$ | -7.98$_{+++}$ | -3.41$_{+++}$ | -4.88$_{+++}$ | -5.13$_{+++}$ | -6.10$_{+++}$ |
| 15 | -13.63 | -10.21$_{+++}$ | -7.48$_{+++}$ | -8.85$_{+++}$ | -8.42$_{+++}$ | -4.10$_{+++}$ | -5.13$_{+++}$ | -5.65$_{+++}$ | -6.07$_{+++}$ |
| 20 | -13.63 | -10.37$_{+++}$ | -8.13$_{+++}$ | -8.90$_{+++}$ | -8.46$_{+++}$ | -4.30$_{+++}$ | -4.84$_{+++}$ | -5.96$_{+++}$ | -6.39$_{+++}$ |
| 25 | -13.62 | -10.08$_{+++}$ | -8.09$_{+++}$ | -8.87$_{+++}$ | -8.53$_{+++}$ | -4.53$_{+++}$ | -4.84$_{+++}$ | -6.32$_{+++}$ | -6.93$_{+++}$ |

Notes: In Panel A, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0 = \mu_{\bar{r}_h^{rw}}$ (*** 1%, ** 5%, * 10%). In Panels B and C, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0$ (*** 1%, ** 5%, * 10%), whereas subscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = \mu_{\bar{r}_h^{rw}}$ (+++ 1%, ++ 5%, + 10%).

**Table A.7**: Rate of return to trading implied volatility: Euro, 2013-2015

A: No Transaction Cost

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.000 | 2.83 | $6.05^{**}$ | $5.57^{**}$ | $5.88^{**}$ | 0.94 | 0.55 | 2.14 | $5.31^{**}$ |
| 2 | 0.000 | -0.39 | 2.94 | $4.36^{**}$ | $5.16^{**}$ | -1.72 | 1.54 | 0.86 | 1.82 |
| 5 | 0.000 | -0.05 | -0.47 | 2.02 | $3.10^{*}$ | -2.44 | -1.54 | 0.74 | 0.94 |
| 10 | 0.000 | 1.58 | 1.50 | 4.59 | 3.73 | -3.22 | -1.87 | 0.80 | -0.05 |
| 15 | 0.000 | 1.39 | $2.91^{*}$ | $5.21^{**}$ | 4.03 | -3.46 | -2.53 | 0.39 | -0.78 |
| 20 | 0.000 | 1.71 | $4.10^{**}$ | $5.94^{**}$ | $5.28^{**}$ | -3.81 | -2.82 | 0.46 | -0.94 |
| 25 | 0.000 | $1.98^{*}$ | $4.87^{***}$ | $6.48^{***}$ | $5.60^{**}$ | -3.76 | -3.00 | 1.01 | -0.55 |

B: 2 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -5.42 | $-0.36_{++}$ | $3.25_{+++}$ | $2.37_{+++}$ | $3.11_{+++}$ | -1.99 | -2.29 | $-0.72_{+}$ | $2.46_{+++}$ |
| 2 | -5.41 | -3.62 | $0.12_{+++}$ | $1.52_{+++}$ | $2.35_{+++}$ | -4.58 | $-1.28_{+}$ | $-1.97_{+}$ | $-0.93_{+}$ |
| 5 | -5.41 | -3.24 | -3.35 | $-1.11_{+}$ | $0.61_{+++}$ | -5.28 | -4.79 | $-2.20_{++}$ | $-1.75_{++}$ |
| 10 | -5.41 | $-1.48_{+++}$ | $-1.25_{+++}$ | $1.74_{+++}$ | $1.07_{+++}$ | -6.13 | -5.14 | $-2.19_{++}$ | -2.69 |
| 15 | -5.42 | $-1.64_{++}$ | $0.22_{+++}$ | $2.49_{+++}$ | $1.31_{+++}$ | -6.42 | -5.79 | -2.68 | -3.41 |
| 20 | -5.42 | $-1.36_{+++}$ | $1.46_{+++}$ | $3.38^{*}_{+++}$ | $2.62_{+++}$ | -6.80 | -6.05 | -2.63 | -3.47 |
| 25 | -5.42 | $-1.10_{+++}$ | $2.25^{**}_{+++}$ | $3.99^{**}_{+++}$ | $2.94^{*}_{+++}$ | -6.81 | -6.19 | $-2.09_{++}$ | -3.12 |

C: 5 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -13.54 | $-5.15_{+++}$ | $-0.97_{+++}$ | $-2.44_{+++}$ | $-1.06_{+++}$ | $-6.39_{+++}$ | $-6.56_{+++}$ | $-5.01_{+++}$ | $-1.80_{+++}$ |
| 2 | -13.54 | $-8.46_{++}$ | $-4.11_{+++}$ | $-2.73_{+++}$ | $-1.86_{+++}$ | $-8.87_{+}$ | $-5.50_{+++}$ | $-6.21_{+++}$ | $-5.07_{+++}$ |
| 5 | -13.54 | $-8.03_{+++}$ | $-7.66_{+++}$ | $-5.81_{+++}$ | $-3.12_{+++}$ | $-9.55_{+}$ | $-9.67_{++}$ | $-6.59_{+++}$ | $-5.79_{+++}$ |
| 10 | -13.54 | $-6.07_{+++}$ | $-5.37_{+++}$ | $-2.55_{+++}$ | $-2.92_{+++}$ | -10.50 | $-10.04_{+}$ | $-6.66_{+++}$ | $-6.64_{+++}$ |
| 15 | -13.54 | $-6.20_{+++}$ | $-3.80_{+++}$ | $-1.59_{+++}$ | $-2.78_{+++}$ | -10.87 | -10.68 | $-7.30_{+++}$ | $-7.34_{+++}$ |
| 20 | -13.54 | $-5.98_{+++}$ | $-2.49_{+++}$ | $-0.46_{+++}$ | $-1.36_{+++}$ | -11.30 | -10.89 | $-7.25_{+++}$ | $-7.26_{+++}$ |
| 25 | -13.55 | $-5.72_{+++}$ | $-1.68_{+++}$ | $0.25_{+++}$ | $-1.05_{+++}$ | -11.38 | -10.99 | $-6.74_{+++}$ | $-6.97_{+++}$ |

Notes: In Panel A, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0 = \mu_{\bar{r}_h^{rw}}$ (*** 1%, ** 5%, * 10%). In Panels B and C, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0$ (*** 1%, ** 5%, * 10%), whereas subscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = \mu_{\bar{r}_h^{rw}}$ (+++ 1%, ++ 5%, + 10%).

**Table A.8**: Rate of return to trading implied volatility: Canadian Dollar, 2010-2012

A: No Transaction Cost

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.00 | 14.48$^{***}$ | 7.71$^{***}$ | 11.18$^{***}$ | 7.82$^{***}$ | 8.92$^{***}$ | 7.17$^{***}$ | 9.83$^{***}$ | 11.98$^{***}$ |
| 2 | 0.00 | 10.05$^{***}$ | 5.93$^{***}$ | 8.83$^{***}$ | 7.02$^{***}$ | 5.15$^{**}$ | 2.35 | 8.04$^{***}$ | 9.15$^{***}$ |
| 5 | 0.00 | 6.99$^{***}$ | 2.05 | 6.01$^{**}$ | 3.12$^{**}$ | 2.89$^{*}$ | -0.60 | 6.17$^{**}$ | 4.36$^{*}$ |
| 10 | 0.00 | 5.30$^{**}$ | 1.00 | 3.37$^{*}$ | -0.51 | 1.13 | -1.35 | 6.39$^{*}$ | 5.45$^{*}$ |
| 15 | 0.00 | 3.44$^{**}$ | 0.60 | 2.16 | -1.02 | 0.91 | -0.94 | 6.49$^{*}$ | 6.16$^{*}$ |
| 20 | 0.00 | 2.28 | 0.03 | 1.88 | -0.93 | 0.32 | -0.37 | 6.32 | 6.06 |
| 25 | 0.00 | 1.49 | 0.36 | 1.65 | -0.54 | 0.47 | -0.16 | 5.93 | 5.91 |

B: 2 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -5.41 | 7.40$^{***}_{+++}$ | 2.62$_{+++}$ | 6.04$^{**}_{+++}$ | 2.62$_{+++}$ | 2.53$_{+++}$ | 1.69$_{+++}$ | 3.65$^{*}_{+++}$ | 5.28$^{**}_{+++}$ |
| 2 | -5.41 | 2.87$_{+++}$ | 1.18$_{+++}$ | 3.47$^{*}_{+++}$ | 1.58$_{+++}$ | -0.72$_{++}$ | -2.91 | 2.16$_{+++}$ | 2.66$_{+++}$ |
| 5 | -5.41 | -1.33$_{+}$ | -4.10 | -1.92$_{+}$ | -4.47 | -2.93 | -5.56 | -0.60$_{++}$ | -1.80$_{+}$ |
| 10 | -5.41 | -3.15 | -4.92 | -3.99 | -7.10 | -4.49 | -6.21 | 0.11$_{++}$ | -0.20$_{+}$ |
| 15 | -5.41 | -4.95 | -5.19 | -4.87 | -6.89 | -4.33 | -5.72 | 0.31$_{+}$ | 0.55$_{++}$ |
| 20 | -5.41 | -6.25 | -5.55 | -4.43 | -6.33 | -4.75 | -5.17 | 0.23$_{+}$ | 0.42$_{++}$ |
| 25 | -5.40 | -6.97 | -5.05 | -4.22 | -5.56 | -4.46 | -4.97 | 0.05$_{+}$ | 0.22$_{++}$ |

C: 5 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -13.53 | -3.20$_{+++}$ | -5.01$_{+++}$ | -1.66$_{+++}$ | -5.19$_{+++}$ | -7.04$_{+++}$ | -6.54$_{+++}$ | -5.63$_{+++}$ | -4.78$_{+++}$ |
| 2 | -13.53 | -7.91$_{++}$ | -5.96$_{+++}$ | -4.57$_{+++}$ | -6.58$_{+++}$ | -9.52$_{+}$ | -10.79 | -6.66$_{+++}$ | -7.07$_{+++}$ |
| 5 | -13.53 | -13.82 | -13.34 | -13.81 | -15.85 | -11.65 | -13.00 | -10.76 | -11.03 |
| 10 | -13.52 | -15.83 | -13.82 | -15.04 | -16.99 | -12.92 | -13.49 | -9.30$_{+}$ | -8.66$_{+}$ |
| 15 | -13.52 | -17.54 | -13.86 | -15.42 | -15.70 | -12.19 | -12.89 | -8.96$_{+}$ | -7.87$_{++}$ |
| 20 | -13.51 | -19.05 | -13.91 | -13.91 | -14.42 | -12.37 | -12.36 | -8.90$_{+}$ | -8.04$_{++}$ |
| 25 | -13.51 | -19.65 | -13.17 | -13.03 | -13.10 | -11.85 | -12.19 | -8.76$_{+}$ | -8.32$_{+}$ |

Notes: In Panel A, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0 = \mu_{\bar{r}_h^{rw}}$ (*** 1%, ** 5%, * 10%). In Panels B and C, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0$ (*** 1%, ** 5%, * 10%), whereas subscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = \mu_{\bar{r}_h^{rw}}$ (+++ 1%, ++ 5%, + 10%).

**Table A.9**: Rate of return to trading implied volatility: Canadian Dollar, 2013-2015

A: No Transaction Cost

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.00 | $3.48^{*}$ | 2.55 | 2.38 | 2.38 | 2.29 | $3.42^{*}$ | 2.78 | $3.38^{*}$ |
| 2 | 0.00 | 2.82 | $4.22^{*}$ | 2.60 | $3.67^{*}$ | 2.23 | $4.30^{*}$ | 3.09 | 3.17 |
| 5 | 0.00 | 2.48 | 3.69 | 3.60 | 3.47 | 3.38 | 3.60 | 3.58 | $3.63^{*}$ |
| 10 | 0.00 | 2.13 | 2.90 | 2.58 | 2.38 | 2.23 | 2.81 | 2.60 | 2.76 |
| 15 | 0.00 | 1.72 | 2.38 | 2.11 | 2.02 | 1.83 | 2.35 | 1.89 | 2.15 |
| 20 | 0.00 | 1.90 | 2.21 | 1.94 | 2.20 | 1.89 | 2.40 | 1.87 | 2.25 |
| 25 | 0.00 | 1.94 | 2.11 | 1.87 | 2.13 | 1.93 | 2.31 | 1.81 | 2.18 |

B: 2 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -5.43 | -2.27 | -3.15 | -3.35 | -3.35 | -3.40 | -2.35 | -2.91 | -2.38 |
| 2 | -5.43 | -2.84 | $-1.58_{+}$ | -3.09 | -2.05 | -3.41 | $-1.46_{+}$ | -2.61 | -2.52 |
| 5 | -5.43 | -3.06 | -2.04 | -2.15 | -2.26 | -2.30 | -2.12 | -2.11 | $-2.05_{+}$ |
| 10 | -5.43 | -3.34 | -2.78 | -3.12 | -3.25 | -3.41 | -2.93 | -3.04 | -2.90 |
| 15 | -5.43 | -3.65 | -3.25 | -3.58 | -3.60 | -3.81 | -3.38 | -3.73 | -3.49 |
| 20 | -5.43 | -3.45 | -3.38 | -3.71 | -3.41 | -3.72 | -3.35 | -3.67 | -3.43 |
| 25 | -5.43 | -3.37 | -3.48 | -3.73 | -3.49 | -3.69 | -3.45 | -3.66 | -3.49 |

C: 5 Basis Points

| *h* \Model | RW | CFL10 | DI10FD | FXVIXC | AVEC | FXVIXL1 | FXVIXL2 | AVEL1 | AVEL2 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -13.57 | -10.92 | -11.70 | -11.97 | -11.95 | -11.94 | -11.01 | -11.46 | -11.02 |
| 2 | -13.57 | -11.33 | -10.28 | -11.63 | -10.63 | -11.86 | -10.10 | -11.16 | -11.07 |
| 5 | -13.57 | -11.37 | -10.64 | -10.80 | -10.85 | -10.83 | -10.71 | -10.65 | -10.57 |
| 10 | -13.57 | -11.55 | -11.30 | -11.68 | -11.70 | -11.88 | -11.53 | -11.52 | -11.39 |
| 15 | -13.57 | -11.72 | -11.71 | -12.12 | -12.04 | -12.29 | -11.99 | -12.15 | -11.94 |
| 20 | -13.57 | -11.49 | -11.76 | -12.19 | -11.84 | -12.13 | -11.98 | -11.99 | -11.96 |
| 25 | -13.57 | -11.36 | -11.88 | -12.14 | -11.93 | -12.13 | -12.08 | -11.88 | -12.01 |

Notes: In Panel A, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0 = \mu_{\bar{r}_h^{rw}}$ (*** 1%, ** 5%, * 10%). In Panels B and C, superscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = 0$ (*** 1%, ** 5%, * 10%), whereas subscripts indicate the significance level of one-tailed tests of whether $H_0$: $\mu_{\bar{r}_h^c} = \mu_{\bar{r}_h^{rw}}$ (+++ 1%, ++ 5%, + 10%).