Forecasting Implied Volatility in Foreign Exchange Markets: A Functional Time Series Approach

Fearghal Kearney[*]   Mark Cummins[†]   Finbarr Murphy[‡]

Abstract

We utilise novel functional time series (FTS) techniques to characterise and forecast implied volatility in foreign exchange markets. In particular, we examine the daily implied volatility curves of FX options, namely; EUR-USD, EUR-GBP, and EUR-JPY. The FTS model is shown to produce both realistic and plausible implied volatility shapes that closely match empirical data during the volatile 2006-2013 period. Furthermore, the FTS model significantly outperforms implied volatility forecasts produced by traditionally employed parametric models. The evaluation is performed under both in-sample and out-of-sample testing frameworks with our findings shown to be robust across various currencies, moneyness segments, contract maturities, forecasting horizons, and out-of-sample window lengths. The economic significance of the results is highlighted through the implementation of a simple trading strategy.

Keywords: Exchange rates; implied volatility; forecasting; functional data analysis; functional time series

JEL Classifications: G10; G15; G17

forthcoming in the European Journal of Finance

[*]Corresponding Author: Queen's Management School, Queen's University Belfast, BT9 5EE, UK. E-mail: F.Kearney@qub.ac.uk.

[†]Dublin City University Business School, Dublin City University, Dublin 9, Ireland. E-mail: Mark.Cummins@dcu.ie.

[‡]Kemmy Business School, University of Limerick, Limerick, Ireland. E-mail: Finbarr.Murphy@ul.ie.

1   Introduction

The Black-Scholes (1973) model assumes that volatility is constant. This assumption, if true, should result in a flat implied volatility curve; the market's expectation of average price volatility for the underlying asset to an option contract between now and its expiry date. Of course in practice, observed implied volatility differs across option contracts, dependent on both moneyness and expiry date. As well as being a transformation of the option price, and a key parameter in many asset pricing formulae, implied volatility is of interest due to its informational content (see Corrado and Miller 2006, Taylor et al. 2010, Muzzioli 2010, and Garvey and Gallagher 2012). Yu et al. (2010) demonstrate this by finding superior results using implied volatility to predict future return volatility of stock index options, when compared to traditional benchmark models in over-the-counter (OTC) and exchange markets. One such OTC market is that of foreign exchange (FX) options. FX is the largest asset class in the world with the Bank for International Settlements reporting that trading levels in FX markets averaged \$5.3 trillion per day.[1] Many stakeholders are exposed to FX risk including banks, speculators, traders, multinational firms, importers, and exporters. Modelling foreign currency cash flows, investment decisions, and hedging strategies, are all greatly dependent on expectations of future FX movements. Our study adds to the existing literature through the novel proposal of a functional time series-based forecasting model to predict the evolution of the implied volatility of three major currency pairs. This is achieved by characterising the implied volatility relationship among option contracts as smooth curves or functions. We not only contribute from an academic perspective, where insights into the dynamics of implied volatility aid our understanding of option markets, but also from a market practitioner perspective, by demonstrating the efficacy of the approach in a trading context.

Compared to previous studies forecasting the volatility of returns, there is a relative paucity of literature predicting the evolution of implied volatility. Examples include Gonçalves and Guidolin (2006), Konstantinidi et al. (2008), Chalamandaris and Tsekrekos (2010), Dunis et al. (2013), Bernales and Guidolin (2014), and Chalamandaris and Tsekrekos (2014). Gonçalves and Guidolin (2006), for instance, analyse S&P 500 Index options, using a Dumas et al. (1998) parametric specification based on moneyness and time to maturity dimensions to characterise the implied volatility surface. They find predictability only on narrow segments of the surface. Konstantinidi et al. (2008), take a different approach, in that they use a number of economic indicators to construct a forecasting model that finds statistically significant predictable patterns in the evolution

---

[1]Bank of International Settlements report available at: http://www.bis.org/publ/rpfx13fx.pdf

of European and U.S. implied volatility indices. Dunis et al. (2013) apply the same economic model to predict the evolution of implied volatility in the EUR-USD exchange rate, a currency pair which we also study. They find that implied volatility is only predictable at short time horizons of up to five hours ahead. The studies that most closely resemble ours are those by Chalamandaris and Tsekrekos (2010, 2011, and 2014). Chalamandaris and Tsekrekos (2010) adopt the methodology of Stock and Watson (2002) to extract latent statistical factors that are then forecast with mixed results. Chalamandaris and Tsekrekos (2011) extend the Gonçalves and Guidolin (2006) framework above by explicitly modelling the term structure of the implied volatility surface along the lines of Diebold and Li (2006). Finally, Chalamandaris and Tsekrekos (2014) compare and contrast the performance of all three proposed approaches, the principal components model of Stock and Watson (2002), and the parametric frameworks offered in Gonçalves and Guidolin (2006) and Chalamandaris and Tsekrekos (2011). The primary conclusion from this strand of literature is that structured parametric forecasting models achieve superior out-of-sample results. We further the literature through an empirical demonstration of statistically significant predictability using an alternative flexible functional approach.

Our functional approach aims to uncover the process underlying a data set and incorporates the shape of the implied volatility smile into its forecast. By making the simple assumption of smoothness in the implied volatility smiles, we can produce infinite dimensional functions that expose additional dynamics missed by traditional multivariate techniques. Functional time series analysis also boasts the advantages of being computationally efficient and of allowing curves to be evaluated on an arbitrarily fine grid.[2] For these reasons despite being in its relative infancy, its popularity in terms of financial applications is growing. Similar to our paper, Benko et al. (2009) assume smoothness of the implied volatility smile in order to conduct functional principal component analysis (FPCA henceforth), analysing similarities in stochastic behaviours between the implied volatility smiles of one and three month option contracts on the German-Swiss equity exchange (EUREX). Muller et al. (2011) study high frequency S&P 500 Index levels, and propose a functional process to characterise volatility trajectories. Their model uncovers patterns in volatility and by combining it with prediction techniques and functional regression, it can be used to predict future volatility levels. Kearney et al. (2015) also use the assumption of smoothness in the smile in order to obtain a measure of implied volatility steepness in oil options using functional techniques. Other financial applications of functional time series include: Kargin and Onatski (2008) on Eu-

_____

[2] These and other advantages of FDA are outlined in Ramsay and Silverman (2005).

rodollar futures; Horvath and Kokoszka (2011) on credit card transactions; Kosiorowski (2014) on predictions of economic time series; Shang (2016) on intraday S&P 500 index returns. In a related field, Antoch et al. (Prchal, De Rosa, and Sarda 2008), Liebl (2013), and Shang (2013) apply FPCA to electricity demand forecasting. In our study we also adopt a FPCA framework, specifically the methodology of Hyndman and Shang (2009). Our study is distinct from prior functional time series studies however, in that we seek to characterise and forecast implied volatility in FX markets. Despite being defined in the functional space, the intuition behind our approach effectively builds on elements from popular implied volatility forecasting models, namely Stock and Watson (2002) and Gonçalves and Guidolin (2006).

General equilibrium models of the implied volatility surface view its shape as being the result of a framework that aggregates investor beliefs. Examples include David and Veronesi (2000), Guidolin and Timmermann (2003), and Garcia et al. (2003). As similarly argued for standard PCA (Chalamandaris and Tsekrekos, 2010) our functional decomposition is consistent with these theories as the latent statistical factors we uncover can be interpreted as proxies for investor uncertainty and learning regarding economic fundamentals that drive the dynamics of the implied volatility surface. In the empirical analysis the first contribution we note is that the flexible functional time series model produces realistic and plausible implied volatility surface shapes that mirror market observations. We show that our proposed functional time series model is competitive and outperforms the popular Gonçalves and Guidolin (2006) and Chalamandaris and Tsekrekos (2011) models with the results being statistically significant in out-of-sample testing. This is consistent across moneyness levels, forecasting horizons, contract maturities, and multiple currencies.

The remainder of the paper is organised as follows. Section 2 provides a background to the functional time series methodology and the forecast evaluation procedure. Section 3 introduces the FX options data set. Section 4 presents and discusses the empirical results, with Section 5 concluding the paper.

2 Methodology

2.1 Functional time series

Functional data analysis provides a functional representation of the process underlying a data set. The functional data methodology has many advantages; it accurately captures implied volatility dynamics (Benko et al. 2009; Kearney et al. 2015), there is no assumed parametric structure, it

is computationally efficient, and it results in a process that can be evaluated on an arbitrarily fine grid. These and other advantages of functional data analysis are outlined in Ramsay and Silverman (2005). A subset of the functional data analysis literature examines models of functional observations that exhibit a temporal relationship; functional time series (FTS). In this paper the functions are defined in the moneyness domain, as we characterise and forecast the evolution of the implied volatility process. The approach applied here effectively blends together elements of two predominant models in the implied volatility modelling literature; namely, the Stock and Watson (2002) principal components approach as adopted by Chalamandaris and Tsekrekos (2010), and the second step of the two-step framework proposed by Gonçalves and Guidolin (2006) and Diebold and Li (2006).

We begin by outlining the process of producing a functional representation of the individual implied volatility smiles (and ultimately a constructed implied volatility surface) at each time point using functional principal component regression from Shang (2013). The approach is similar in spirit to the Stock and Watson (2002) principal component analysis (PCA) decomposition that is used for forecasting purposes by Chalamandaris and Tsekrekos (2010). The advantages of PCA is that it reduces the dimensionality of the data whilst preserving the maximum amount of dataset information. The most important distinction between this and our functional time series approach is the concept of smoothness. In line with Benko et al. (2009) and Kearney et al. (2015) we make the assumption that the implied volatility smile process is both continuous and a smooth curve.[3] It allows us to view the discrete implied volatility data as a infinite dimensional function upon which we can apply dimension reduction functional principal component regression (FPCR). The implementation requires slight modification from the traditional multivariate PCA method of Stock and Watson (2002) as we will outline.

Firstly, we need to define the problem more explicitly. We have discrete option implied volatility data observed at a daily frequency for each day t, $x_t(m)$. The continuous domain, m, is that of the moneyness level (in terms of delta) of the option contract. Applying our assumption of smoothness we seek to uncover a continuous function without jumps, denoted $\tilde{x}_t(m)$, that characterise (with error) the discretised daily implied volatility smile dynamics. In practice, we observe $\{m_k, x_t(m_k)\}$ for $t = 1, 2, ..., n$ and $k = 1, 2, ..., q$, from which we extract a smooth function $\tilde{x}_t(m)$, given by

$$x_t(m_k) = \tilde{x}_t(m_k) + \sigma_t(m_k)\varepsilon_{t,k}, \qquad (1)$$

where $\varepsilon_{t,k}$ is an independent and identically distributed standard normal random variable, $\sigma_t(m_k)$ allows the amount of noise to vary with m_k, and $\{m_1, m_2, ..., m_q\}$ is the set of discrete delta values.

---

[3]We only assume smoothness across the implied volatility smile as prior functional data literature has yet to establish smoothness of the implied volatility surface term structure.

Given a set of functional data, denoted $\tilde{\mathbf{x}}(\mathbf{m})$, where $\tilde{\mathbf{x}}(\mathbf{m}) = [\tilde{x}_1(m), \tilde{x}_2(m), ..., \tilde{x}_t(m)]^{\top}$, we seek to produce a realistic reduced form parsimonious model with orthogonal regressors and uncorrelated regression coefficients using FPCR.

We will now introduce functional principal component analysis for the reader. FPCA commences with the search for weight functions, $\phi_j(m)$, that correspond to probe or functional principal component scores, $\beta_{t,j}$, with the highest possible levels of variation. To ensure that each new principal component function captures a distinct mode of variation, they are required to be orthogonal to those computed previously:

$$\int \phi_r(m)\phi_l(m)dm = 0 \quad r = 1, ..., l - 1.$$

In our sample we observe $n = 2055$ realisations of $\tilde{x}(m)$ evaluated on a compact interval of delta, $m \in [5, 95]$, for each contract maturity. Therefore, at a sample level, the functional principal component decomposition can be represented as:

$$\tilde{x}_t(m) = \bar{\tilde{x}}_t(m) + \sum_{j=1}^{J} \hat{\beta}_{t,j}\hat{\phi}_j(m) + \hat{\varepsilon}_t(m), \qquad (2)$$

where $\bar{\tilde{x}}_t(m) = \frac{1}{n}\sum_{t=1}^{n} \tilde{x}_t(m)$ is the estimated mean function, $\hat{\phi}_j(m)$ is an estimate of the jth orthonormal eigenfunction of the empirical covariance operator, defined as:

$$\hat{\Gamma}(m) = \frac{1}{n}\sum_{t=1}^{n} [\tilde{x}_t(m) - \bar{\tilde{x}}_t(m)][\tilde{x}_t(m) - \bar{\tilde{x}}_t(m)].$$

The coefficient $\beta_{t,j}$ is the jth principal component score for day t.[4] Based on Chalamandaris and Tsekrekos (2010, 2014) who decompose the implied volatility surface into three principal components we specify $J = 3$.

The forecasting portion of the functional methodology follows the second step of the Gonçalves and Guidolin (2006) two-step forecasting method. We explicitly account for the dependency structure between sequential functional observations by modelling the evolution of the coefficients as a univariate time series. Specifically, we condition on the observed data, $\tilde{\mathbf{x}}(\mathbf{m}) = [\tilde{x}_1(m), \tilde{x}_2(m), ..., \tilde{x}_t(m)]^{\top}$, and the fixed functional principal components $\mathbf{\Phi}(\mathbf{m}) = \left[\hat{\phi}_{\mathbf{1}}(\mathbf{m}), \hat{\phi}_{\mathbf{2}}(\mathbf{m}), ..., \hat{\phi}_{\mathbf{J}}(\mathbf{m})\right]^{\top}$, to produce the h step ahead forecasts of $x_{n+h}(m)$:

$$\hat{x}_{n+h}(m) = E[x_{n+h}(m) \mid \tilde{\mathbf{x}}(\mathbf{m}), \mathbf{\Phi}(\mathbf{m})] = \bar{\tilde{x}}_t(m) + \sum_{j=1}^{J} \hat{\beta}_{n+h|n,j}\hat{\phi}_j(m),$$

---

[4] Further technical detail of FPCR implementation are provided in Hyndman and Shang (2009).

where $\hat{\beta}_{n+h|n,j}$ denotes the h step ahead forecasts of $\beta_{n+h,j}$ using an autoregressive univariate model of order h.

2.2 Forecast evaluation

We assess the forecast performance of the FTS model using the following measures:

1. Mean absolute error (MAE) is the average of the absolute differences between the forecast, $\hat{x}_{t+1}(m_k)$, and the corresponding observation, $x_{t+1}(m_k)$. It measures the average error magnitude in the forecasts, regardless of error direction and serves to aggregate the errors into a single measure of predictive power.

$$MAE = \frac{1}{n}\sum_{i=1}^{n-1}|x_{t+1}(m_k) - \hat{x}_{t+1}(m_k)|,$$

where $x_{t+1}(m_k)$ are the observed values and $\hat{x}_{t+1}(m_k)$ are the values predicted from the model.

2. Root mean squared error (RMSE) is a measure of the difference between values predicted by a model and realised values. The RMSE is defined as the square root of the mean squared error, and again serves to aggregate the errors into a single measure of predictive power.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n-1}(x_{t+1}(m_k) - \hat{x}_{t+1}(m_k))^2}{n}},$$

where $x_{t+1}(m_k)$ are the observed values and $\hat{x}_{t+1}(m_k)$ are the values predicted from the model.

3. Mean mixed error (MME) is an asymmetric loss function. MME(U) penalises under-predictions more heavily, while MME(O) penalises over-predictions more heavily. This is very important for investors in option markets, as an under (over)-prediction of implied volatility is more likely to be of greater concern to a seller (buyer) than a buyer (seller). The measure has been employed previously in studies evaluating volatility forecasting techniques such as Brailsford and Faff (1996) and Fuertes et al. (2009).

$$MME(U) = \frac{1}{n}\left[\sum_{t=t_1^O}^{t_N^O}|x_{t+1}(m_k) - \hat{x}_{t+1}(m_k)| + \sum_{t=t_1^U}^{t_N^U}\sqrt{|x_{t+1}(m_k) - \hat{x}_{t+1}(m_k)|}\right]$$

and

$$MME(O) = \frac{1}{n}\left[\sum_{t=t_1^O}^{t_N^O} \sqrt{|x_{t+1}(m_k) - \hat{x}_{t+1}(m_k)|} + \sum_{t=t_1^U}^{t_N^U} |x_{t+1}(m_k) - \hat{x}_{t+1}(m_k)|\right],$$

where t_N^U is the number of under-predictions and t_N^O is the number of over-predictions. $t_1^O,...,t_N^O$ represent the indices of the over-predictions, and $t_1^U,...,t_N^U$ represent the indices of the under-predictions.

4. The mean correct predictor of direction of change (MCPDC) is the percentage of predictions for which the forecast, $\hat{x}_{t+1}(m_k)$, has the same sign as the corresponding observation, $x_{t+1}(m_k)$. MCPDC measures how well the model can forecast the direction of movement, regardless of error magnitude. It is also employed in Bernales and Guidolin (2014).[5]

The out-of-sample performance of the FTS prediction is benchmarked against the two leading forecasting models used in the literature, namely, Gonçalves and Guidolin (2006) and Chalamandaris and Tsekrekos (2011). Gonçalves and Guidolin (2006) model daily implied volatility surfaces using Dumas et al. (1998) parametric specifications based on moneyness and time to maturity. Forecasts of this fitted model are produced by assuming the coefficients evolve according to standard time series techniques. A predicted implied volatility surface is then reconstructed using these forecasted coefficients. Chalamandaris and Tsekrekos (2011) also adopt the Dumas et al. (1998) parametric structure but extend their work by stating that the linear approximation of maturity is not sufficient, proposing the use of Nelson-Siegel term structure factors in the spirit of Diebold and Li (2006) to produce fitted implied volatility surfaces. Again, as in Gonçalves and Guidolin (2006), a two-step framework is proposed where forecasts are produced by modelling the fitted coefficients from the fitting stage. Given the success in both Gonçalves and Guidolin (2006) and Chalamandaris and Tsekrekos (2011) of modelling the coefficients using a standard univariate autoregressive model we adopt same. The reader is directed to Chalamandaris and Tsekrekos (2014) for further technical implementation details for both models. Despite showing promising results in terms of predictable segments of the implied volatility surface, forecasts from previous models (including Gonçalves and Guidolin 2006; Konstantinidi et al. 2008; Chalamandaris and Tsekrekos 2010, 2011, and 2014; and Bernales and Guidolin 2014) have been unable to consistently outperform univariate benchmarks. For this reason we qualify our functional time series results by implementing the Diebold-Mariano

---

[5]Bekiros and Georgoutsos (2008) also incorporate directional metrics when studying the dynamics of the CBOE volatility index.

(1995) test of superior forecasting ability versus an autoregressive benchmark model also.

To control for sensitivity to specific out-of-sample periods, various window lengths are tested: 100 day (out-of-sample: July 2013 to November 2013), 200 day (out-of-sample: February 2013 to November 2013), 500 day (out-of-sample: December 2011 to November 2013), and 1000 day (out-of-sample: January 2010 to November 2013). The out-of-sample forecast, between the end of the in-sample period and November 2013, are obtained using a recursive scheme. Each day an additional observation is added to an expanding training window and the models are re-estimated. This is in line with Chalamandaris and Tsekrekos (2010) who adopt a recursive 1-day strategy scheme. Konstantinidi et al. (2008) and Gonçalves and Guidolin (2006), also implement out-of-sample recursive schemes by expanding the training window size at 100-day intervals. We choose to expand the training set and re-estimate the model at each time step, daily, to incorporate all available up-to-date information into our prediction. This approach more accurately simulates the action likely to be taken by a market practitioner who seeks to predict the following day's movement.

3   Data description

The data set comprises, at-the-money, risk reversal, and butterfly composition implied volatility quotes for the Euro/United States Dollar (EUR-USD), Euro/British Pound (EUR-GBP), and Euro/Japanese Yen (EUR-JPY) currency pairs obtained from Bloomberg. These four currencies represent almost 78% of total global foreign exchange market turnover[6] and are also considered by Chalamandaris and Tsekrekos (2011). Focus on these heavily traded currency pairs minimise issues around data quality (i.e., stale and out-of-context quotes). They constitute developed pairs whereby option contracts are the main avenue through which investors exploit the interest rate differentials between the different countries. The use of a contributory data vendor such as Bloomberg, mitigates the idiosyncratic effect specific to individual market participants providing quotes. This issue is cited by Chalamandaris and Tsekrekos (2014), with Bloomberg being used to validate their proprietary J.P. Morgan data set. Through the use of this J.P. Morgan database, Chalamandaris and Tsekrekos (2014) find that implied volatility is more predictable for very liquid currency pairs, citing EUR-USD as an example. EUR-USD is also the sole focus of the study by Dunis et al. (2013). Using only option expiry dates of less than a year; constant option maturities of one, three, six,

---

[6]Bank of International Settlements report available at: http://www.bis.org/publ/rpfx13fx.pdf

and nine months showcases the stark improvement of exploiting the information along the implied volatility smile using our functional techniques. This echoes Chalamandaris and Tsekrekos (2011) who state that a linear approximation of the implied volatility term structure is unproblematic when modelling and forecasting future implied volatility movements for options with expiries of less than a year. It also mitigates any remaining concerns around illiquidity in the less actively traded long term maturity option contracts. Delta values of 5, 10, 15, 25, 35, 50, 65, 75, 85, 90, 95 are constructed from the at-the-money, risk reversal, and butterfly implied volatility quotes using the Black-Scholes (1973) and its Garman and Kohlagen (1983) option pricing extension. Log changes in implied volatility are calculated for the January 2006 to November 2013 period. As in Chalamandaris and Tsekrekos (2011), we limit our forecast prediction to the surfaces with the highest levels of liquidity. The most liquid contracts are delta values of 10, 25, 50, 75, and 90. It is for this reason that our forecasts concentrate on these particular contracts.

4 Empirical results

This section presents the results of modelling the evolution of implied volatility using the functional time series model (FTS, henceforth), the Chalamandaris and Tsekrekos (2011) model (CT11, henceforth), and the Gonçalves and Guidolin (2006) model (GG06, henceforth) for the entire sample, January 2006 to November 2013. Firstly, the models are fitted in-sample to ascertain how well they capture the empirical dynamics of the implied volatility surfaces for each currency during the period. The resultant three-dimensional surfaces are plotted and used as an exploratory tool to provide an intuitive graphical demonstration of which models best fit the underlying dataset. The focus of the paper is on analysing the predictive capacity of the models however. To this end, the in-sample forecasts are initially compared and evaluated using the measures outlined in Section 4.2. Finally, we move to an out-of-sample testing environment where the predictions are analysed across various contract maturities and moneyness levels. To formally establish if the functional model outperforms the benchmark models in terms of forecast accuracy, the Diebold-Mariano (1995) test for superior predictive ability is set out in Section 4.3.

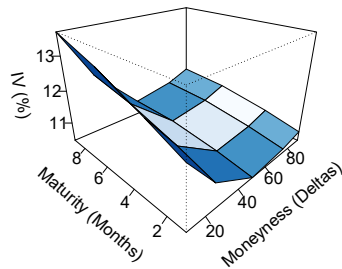Figure 1: EUR-USD average observed and fitted implied volatility
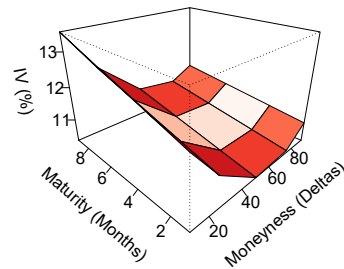
EURUSD IVS (2006−01−04 to 2013−11−20)
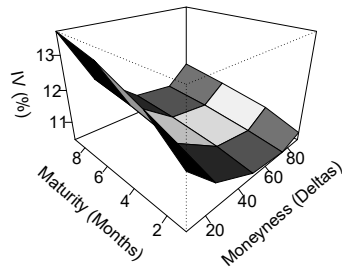
EURUSD FTS IVS (2006−01−04 to 2013−11−20)
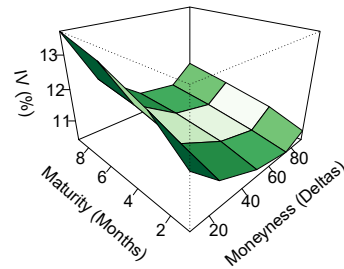
EURUSD CT11 IVS (2006−01−04 to 2013−11−20)

EURUSD GG06 IVS (2006−01−04 to 2013−11−20)

The dependency value uncovered when fitting the fully functional model to the EUR-USD implied volatility data set for the January 2006 to November 2013 period. Option maturity=1 month. "IV" is used as an abbreviation for implied volatility.

Figure 2: EUR-GBP average observed and fitted implied volatility

EURGBP IVS (2006−01−04 to 2013−11−20)

EURGBP FTS IVS (2006−01−04 to 2013−11−20)

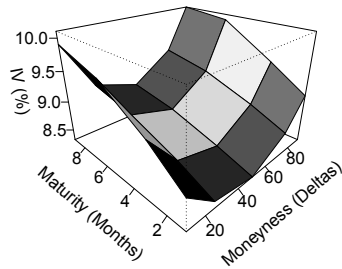
EURGBP CT11 IVS (2006−01−04 to 2013−11−20)

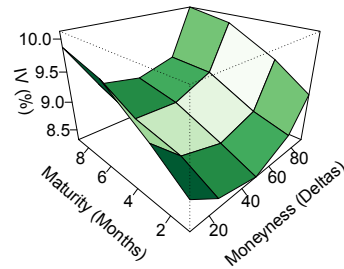**EURGBP GG06 IVS (2006−01−04 to 2013−11−20)**

The dependency value uncovered when fitting the fully functional model to the EUR-USD implied volatility data set for the January 2006 to November 2013 period. Option maturity=1 month. "IV" is used as an abbreviation for implied volatility.

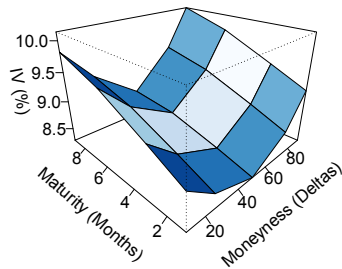Figure 3: EUR-JPY average observed and fitted implied volatility
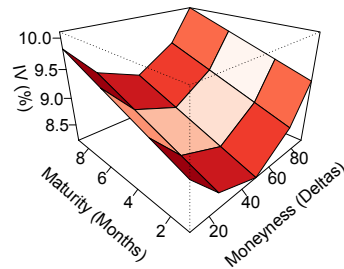
EURJPY IVS (2006−01−04 to 2013−11−20)

EURJPY FTS IVS (2006−01−04 to 2013−11−20)

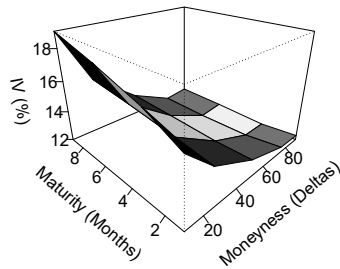**EURJPY CT11 IVS (2006−01−04 to 2013−11−20)**

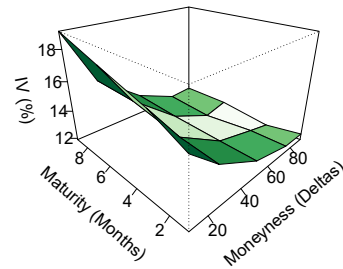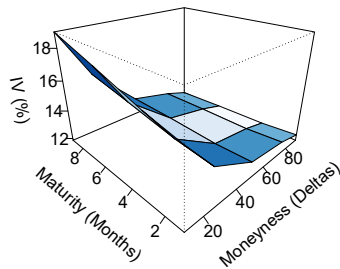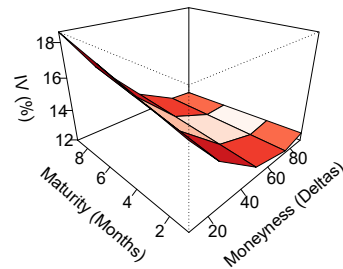**EURJPY GG06 IVS (2006−01−04 to 2013−11−20)**

The dependency value uncovered when fitting the fully functional model to the EUR-USD implied volatility data set for the January 2006 to November 2013 period. Option maturity=1 month. "IV" is used as an abbreviation for implied volatility.

4.1   Model fitting

FTS (equation 2), CT11, and GG06 models are fitted to the underlying implied volatility data outlined in Section 3, for each day over the full sample of January 2006 to November 2013. Implied volatility surface graphs averaged over the period are then produced and plotted in Figures 1, 2, and 3 to facilitate a cross comparison of models.

Firstly, turning our attention to the EUR-USD plots in Figure 1, we analyse the empirically observed surface plot in the the top left panel. The underlying pair traded in the $1.15 to $1.60 range over the period with a high being hit in 2008 and a gradual decrease in the strength of the Euro after that point. The elevated level of implied volatility, ranging from 10.5% to 13.5% highlights investor uncertainty about economic fundamentals and related future currency movements. Analysing the shape, the predominant feature is the asymmetric implied volatility smile. The negative skew, whereby out of the money (OTM) Euro puts (conversely Dollar calls) demonstrate a much higher implied volatility than OTM Euro calls (conversely Dollar puts), conveying investor beliefs that the Euro is set to depreciate versus the Dollar. Scanning the implied volatility surface term structure we can see that the market's expectation of future realised volatility is even higher for medium term maturity options, between six and nine months, than for short term options. That being said, increases in implied volatility occurs most sharply between one and three months maturity, whereas the rate of increase between higher maturity levels is not as pronounced.

Now that we have examined the dynamics at play in the observed implied volatility data we analyse how well the various models perform in capturing these dynamics. Firstly, we can see that the skew in the implied volatility smile is most closely captured by the functional time series model in the top right panel of Figure 1. The lowest point in the implied volatility smile of the observed data is at a delta value of greater than 60. The functional time series model mirrors this whereas both CT11 and GG06 show the 50 delta region as being the lowest point, more in line with traditional symmetric implied volatility smiles. Comparing the three models in terms of modelling the term structure shows more subtle differences than across the moneyness domain. GG06 use a linear approximation which, graphically at least, does not appear to be sufficient. CT11 demonstrates the advantages of a non-linear fit of the term structure, yet it does not capture the shape as accurately as the FTS model.

Secondly, we focus on the EUR-GBP pair. The underlying pair traded in the £0.65 to £0.95 range over the sample, however the 2009 to 2013 period is relatively benign, within which Sterling appreciates slightly reversing some of the depreciation seen during the crisis period. The EUR-GBP implied volatility surface is depicted in Figure 2. Lower implied volatility levels are observed, with values ranging from 8.5% to 10%. The observed plot does not exhibit the same negative skew as exhibited by the EUR-USD surface. The smile observed for EUR-GBP is almost symmetric with the turning point being located in the ATM 50 delta region. Increased uncertainty for longer dated maturities is also observed, as in the case of the EUR-USD. Increasing contract maturity we observe a sharp increase in implied volatility up to six months maturity with a levelling off of increases between the six and nine month maturities. The question we now ask is how well do our models capture this implied volatility shape?

The short answer is, rather well. Across the moneyness dimension all three models perform well, capturing the turning point of the smile and the extreme OTM implied volatilities accurately. This more traditional shape appears to suit the parametric CT11 and GG06 models, however the FTS still matches their accuracy (at least in this graphical representation). Analysing the term structure, GG06 is again let down by approximating only a linear relationship between options of different maturities, whilst CT11 shows limited deviation from the straight line in this case. The FTS model on the other hand successfully captures the term structure relationship, demonstrating a steep incline up to and including six month maturity and less steep thereafter.

Lastly, we analyse the EUR-JPY pair. Over this period the underlying currency rate for EUR-JPY trades in the 95 Yen to 170 Yen range and exhibits large swings throughout. The implied volatility surface is shown in Figure 3.. The observed implied volatility data demonstrates a strong negative skew. This indicates that market expectations over the period are that the Yen will appreciate, in both short and medium term forecasts. The absolute level of implied volatility for the pair is very high, with values ranging from 12% to 18%. The empirical term structure is relatively flat showing that the expectation of JPY appreciation dominates option pricing, having broadly similar affects for all option contracts. All three models perform well in terms of fitting to this data. One criticism of the CT11 and GG06 models in comparison to the FTS model is that their smile turning points are at implied volatility levels that are low when compared to the observed curve whose lowest implied volatility level is in the 80 delta region. Despite the term structure not being a primary driver of the shape of the implied volatility surface for this pair, the simplest parametric model, GG06, thrives in this environment due to it's linear approximation of the maturity.

The aim of these figures is to provide an intuitive and accessible representation for a general

Table 1: In-sample performance measures

Model	$\overline{RMSE}$	$\overline{MCPDC}$	$\overline{RMSE}$	$\overline{MCPDC}$	$\overline{RMSE}$	$\overline{MCPDC}$	$\overline{RMSE}$	$\overline{MCPDC}$
	1 Month		3 Month		6 Month		9 Month	
				USD				
FTS	0.0377	0.5158	0.0284	0.4925	0.0230	0.4658	0.0223	0.4220
CT11	0.0422	0.4805	0.0753	0.4321	0.1253	0.4406	0.1464	0.4547
GG06	0.0388	0.4591	0.0299	0.4771	0.0241	0.4915	0.0229	0.4767
				GBP				
FTS	0.0341	0.5164	0.0267	0.4506	0.0230	0.4531	0.0215	0.3767
CT11	0.0407	0.4249	0.0765	0.3664	0.1263	0.3745	0.1476	0.3387
GG06	0.0419	0.4416	0.0351	0.4135	0.0317	0.4226	0.0298	0.3502
				JPY				
FTS	0.0532	0.4825	0.0400	0.4775	0.0322	0.4607	0.0289	0.4233
CT11	0.0568	0.5041	0.0821	0.4577	0.1313	0.4481	0.1519	0.4539
GG06	0.0657	0.4517	0.0538	0.4659	0.0467	0.4589	0.0439	0.4321

$\overline{RMSE}$ and $\overline{MCPDC}$ represent RMSE and MCPDC performance metrics averaged across the delta values available for a one-day ahead forecast of EUR-USD, EUR-GBP, and EUR-JPY implied volatility data. The period covered is the in-sample of January 2006 to the January 2006 to January 2010 period. The values are provided for the Functional Time Series (FTS), CT11 model, and the GG06 model.. Option maturities of 1, 3, 6, and 9 months are given separately.

audience of some of the strengths and weaknesses of each model adopted. In summary, the FTS model performs well in modelling the dynamics across the moneyness for each of the three currencies. Both the parametric models from the literature, the GG06 and CT11 underperform, in particular in the extreme cases of stressed market environments, however, they perform relatively well in modelling the implied volatility surface of the less volatile EUR-GBP pair. Notwithstanding the fact that both parametric models simultaneously exploit information from the full implied volatility surface, the improvement in term structure modelling exhibited by the FTS model reflects its ability to more accurately extract information from the isolated smiles separately; indirectly leading to a well-represented implied volatility surface reconstruction.

4.2   In-sample predictions

The promising fitting results for the FTS model across each of the currencies lead us first of all, to test if it can be employed to establish an intertemporal dependency across implied volatility surfaces on subsequent days. We go about this in two stages, firstly, we test the predictions produced by our models in-sample, and secondly, we split the sample into both training and out-of-sample segments to see how well each performs. We begin by looking at the performance metrics produced from comparing our forecasts with the actual values.

For brevity, Table 1 lists RMSE and MCPDC values averaged across the implied volatility smile. Metrics are calculated separately for each of the contract maturities. For one month EUR-USD and

EUR-GBP options the FTS most accurately forecasts based on both direction of change and RMSE metrics. It correctly identifies the direction of the subsequent days change in implied volatility in over 51.50% of cases for these two currencies. The results for the longer dated maturity options also demonstrate the advantages of adopting the FTS forecasts. The average RMSE figures seen here are consistently lower than those calculated for the one month maturity, and more importantly in comparison with the other parametric models from the literature the FTS forecasts exhibit a lower average error. It should be noted however that the GG06 model exhibits strong RMSE metrics for both six and nine month EUR-USD options, with values of 0.0241 and 0.0223 respectively (versus 0.0230 and 0.0223 for the FTS framework). Success in terms of minimising error does not directly map into accurately predicting directional change, as can be see by the FTS model being out performed by the CT11 model in terms of MCPDC for two nine month maturity options, namely USD (0.4547 versus 0.4220) and JPY (0.4539 versus 0.4233). In both cases the GG06 model also outperforms the FTS framework in terms of directional forecasting accuracy. The CT11 model however exhibits average RMSE values that actually increase significantly for longer maturity options, an indication that the linear approximation of the term structure is sufficient over this volatile in-sample period of January 2006 to January 2010. The Diebold and Li (2006) inspired adaptation of the Nelson-Siegel factors do not appear to be as effective in the environment as they might be during a more benign period.

4.3   Out-of-sample forecast evaluation

It is established in the previous section that the FTS model provides a good in-sample fit for modelling the evolution of implied volatility. We now turn our attention to out-of-sample forecasting. A summary of the out-of-sample forecast measures calculated for at-the-money implied volatility under a recursive parameter estimation scheme and 500 day out-of-sample window length are presented in Table 2. To assess performance across the implied volatility smile, measures for one month maturity for other delta values are given in Table 3.[7] The results from both tables give clear indications that the FTS model outperforms the traditionally used GG06 and CT11 models in forecasting implied volatility out-of-sample over the December 2011 to November 2013 period where a one-day ahead forecasting horizon is implemented.

Firstly, we analyse Table 2. When conducting a straight comparison of metrics, the FTS model

---

[7]Other out-of-sample window periods, of 100, 200 and 1000 days, are utilised with similar results obtained.

Table 2: ATM out-of-sample forecast performance measures

Model	RMSE	MAE	MCPDC	MME(O)	MME(U)	RMSE	MAE	MCPDC	MME(O)	MME(U)
					USD					
			1 Month					3 Month		
FTS	0.0323	0.0251	0.5340	0.0803	0.0895	0.0243	0.0183	0.4960	0.0641	0.0762
CT11	0.0404	0.0305	0.4820	0.0771	0.1121	0.0717	0.0605	0.4420	0.1736	0.1153
GG06	0.0353	0.0273	0.4860	0.0845	0.0925	0.0275	0.0212	0.4640	0.0732	0.0810
			6 Month					9 Month		
FTS	0.0193	0.0144	0.514	0.0565	0.0661	0.0176	0.0132	0.4440	0.0520	0.0650
CT11	0.1216	0.0959	0.4580	0.2251	0.1456	0.1441	0.1094	0.4800	0.2462	0.1470
GG06	0.0225	0.0171	0.5080	0.0639	0.0723	0.0200	0.0152	0.5220	0.0594	0.0681
					GBP					
			1 Month					3 Month		
FTS	0.0323	0.0246	0.4840	0.0791	0.0881	0.0218	0.0165	0.4940	0.0616	0.0712
CT11	0.0393	0.0303	0.5060	0.0763	0.1125	0.0666	0.0567	0.4480	0.1657	0.1147
GG06	0.0372	0.0291	0.4640	0.0917	0.0940	0.0283	0.0220	0.4580	0.0757	0.0816
			6 Month					9 Month		
FTS	0.0172	0.0128	0.466	0.0525	0.0626	0.0152	0.0113	0.4900	0.0489	0.0581
CT11	0.1122	0.0886	0.4600	0.2140	0.1408	0.1327	0.1006	0.4860	0.2322	0.1409
GG06	0.0246	0.0189	0.4420	0.0689	0.0750	0.0227	0.0175	0.4660	0.0668	0.0711
					JPY					
			1 Month					3 Month		
FTS	0.036	0.0275	0.4860	0.0863	0.0912	0.0253	0.0192	0.4600	0.0675	0.0767
CT11	0.0418	0.0321	0.5080	0.0823	0.1126	0.0808	0.0679	0.4580	0.1858	0.1247
GG06	0.0454	0.0355	0.5140	0.1044	0.1030	0.0369	0.0291	0.4940	0.0925	0.0928
			6 Month					9 Month		
FTS	0.0194	0.0145	0.4840	0.0578	0.0649	0.0162	0.012	0.4820	0.0519	0.0581
CT11	0.1382	0.1072	0.4860	0.2378	0.1568	0.1637	0.1224	0.5120	0.2589	0.1584
GG06	0.0326	0.0252	0.5040	0.0859	0.0843	0.0301	0.0234	0.4860	0.0830	0.0809

One-day ahead forecasting performance calculated for at-the-money EUR-USD, EUR-GBP, and EUR-JPY implied volatility data under a recursive out-of-sample parameter estimation scheme and a 500 day out-of-sample window length over the December 2011 to November 2013 period. The values are provided for the Functional Time Series (FTS), CT11 and the GG06 model. Option maturities of 1, 3, 6, and 9 months are given separately.

Table 3: One month out-of-sample forecast performance measures

Model	RMSE	MAE	MCPDC	MME(O)	MME(U)	RMSE	MAE	MCPDC	MME(O)	MME(U)
					USD					
			10 Delta					25 Delta		
FTS	0.0356	0.0275	0.518	0.0842	0.0951	0.0336	0.026	0.51	0.0819	0.0921
CT11	0.0423	0.032	0.51	0.0817	0.1124	0.041	0.031	0.48	0.0785	0.1128
GG06	0.0358	0.0274	0.504	0.0857	0.0927	0.0339	0.026	0.494	0.0827	0.0906
			75 Delta					90 Delta		
FTS	0.0316	0.0244	0.554	0.0794	0.0874	0.0326	0.0254	0.54	0.0819	0.0898
CT11	0.0394	0.0297	0.484	0.0766	0.1094	0.0396	0.0301	0.518	0.0802	0.1076
GG06	0.0438	0.0344	0.486	0.099	0.1058	0.0545	0.0431	0.474	0.1152	0.1189
					GBP					
			10 Delta					25 Delta		
FTS	0.0347	0.0270	0.5760	0.0861	0.0913	0.0326	0.0250	0.5280	0.0794	0.0894
CT11	0.0402	0.0313	0.5300	0.0809	0.1124	0.0392	0.0302	0.5060	0.0765	0.1122
GG06	0.0358	0.0281	0.4580	0.0890	0.0927	0.0334	0.0255	0.4960	0.0819	0.0888
			75 Delta					90 Delta		
FTS	0.0326	0.0250	0.5400	0.0801	0.0888	0.0360	0.0282	0.5260	0.0904	0.0916
CT11	0.0390	0.0302	0.5240	0.0771	0.1117	0.0403	0.0319	0.5200	0.0835	0.1126
GG06	0.0515	0.0400	0.4960	0.1099	0.1128	0.0701	0.0534	0.4620	0.1317	0.1312
					JPY					
			10 Delta					25 Delta		
FTS	0.0383	0.0292	0.5180	0.0889	0.0954	0.0367	0.0277	0.4820	0.0851	0.0926
CT11	0.0434	0.0334	0.5260	0.0889	0.1111	0.0422	0.0321	0.5100	0.0834	0.1113
GG06	0.0385	0.0294	0.4660	0.0903	0.0950	0.0371	0.0280	0.5080	0.0868	0.0925
			75 Delta					90 Delta		
FTS	0.0369	0.0284	0.5120	0.0883	0.0936	0.0405	0.0309	0.5760	0.0941	0.0961
CT11	0.0420	0.0322	0.5260	0.0831	0.1126	0.0437	0.0338	0.5380	0.0887	0.1126
GG06	0.0747	0.0565	0.5100	0.1405	0.1308	0.1043	0.0778	0.4940	0.1707	0.1579

This table provides one-day ahead forecasting performance calculated for one month maturity EUR-USD, EUR-GBP, and EUR-JPY implied volatility data under a recursive out-of-sample parameter estimation scheme and a 500 day out-of-sample window length over the December 2011 to November 2013 period. The values are provided for the Functional Time Series (FTS), CT11 model, and the GG06 model. Delta values of 10, 25, 75, 90 are given separately.

outperforms in terms of both RMSE and MAE across all currencies and maturity lengths. The MCPDC results are also positive, specifying that the FTS correctly predicts the direction of implied volatility change up to 53.40% of the time as seen for one month maturity EUR-USD options. In the previous in-sample section however, we established that minimising the prediction error does not map directly into a particular model displaying outperformance in terms of directional capabilities. This is seen most prominently for the Japanese Yen, as the FTS model fails to beat the MCPDC metric of either CT11 or GG06 for any JPY contract maturity.[8] Directional predictability performance for the EUR-USD and EUR-GBP is more in line with the signal given from the strong RMSE and MAE results however, with near systematic FTS outperformance, apart from MCPDC for nine month maturity EUR-USD, (FTS 0.4440 versus GG06 0.4800 and CT11 0.5220) and one month maturity EUR-GBP (FTS 0.4840 versus GG06 0.4640 and CT11 0.5060). This strong MCPDC result give an indication of the potential profitability of implementing a trading strategy based on the functional time series model, primarily for the EUR-USD and EUR-GBP currencies. The asymmetric mean mixed error loss functions give an indication of which models systematically under- and over- predict implied volatility changes. The closer the MME(U) and MME(O) values for a given model, the lower the level of systematic under- or over- prediction. The MME(O) and MME(U) results presented in Table 2 indicate that the FTS model has a slight tendency to under-predict future implied volatility change. The one month maturity EUR-USD MME(O) and MME(U) values of 0.0803 and 0.0895, respectively, are quite close however, indicating that any bias is minor. This tendency to under-predict ATM implied volatility is seen systematically across all currencies and contract maturities tested. The CT11 shows an interesting dynamic also, in that across all three currencies it under-predicts the ATM implied volatility over the period for the one month contract maturity, yet it over-predicts the three, six, and nine month contract maturities. The GG06 model shows ATM under-predictions for both EUR-USD and EUR-GBP currencies, however for EUR-JPY ATM implied volatilities it show remarkably unbiased predictions, with MME(O) and MME(U) asymmetric loss figures of 0.0925 and 0.0928, respectively, for three month maturity options.

According to Chalamandaris and Tsekrekos (2010), non-uniform trading causes segments of the implied volatility surface to adjust to information at different rates. For this reason we now turn our attention to predicting implied volatility across the smile, with performance metrics for the non-ATM options; 10 delta, 25 delta, 75 delta, and 90 delta, of the one month contract maturity

---

[8]In results available upon request, despite not outperforming the CT11 and GG06 models, the FTS model does out perform a univariate autoregressive benchmark in terms of MCPDC for all EUR-JPY out-of-sample implied volatility contract maturities.

being shown in Table 3.[9] Firstly, we compare and contrast RMSE and MAE metrics across the three models. Mirroring the outperformance observed for the ATM case we find that the FTS model systematically outperforms across all three currencies. Broadly speaking, FTS exhibits more striking outperformance versus the two parametric models for higher values of delta (OTM EUR calls versus OTM EUR puts). One such example of this is the 10 Delta one month EUR-USD contract (RMSE values of 0.0356, 0.0423, and 0.0358 for FTS, CT11, and GG06 respectively) versus the 90 Delta one month EUR-USD contract (RMSE values of 0.0326, 0.0396, and 0.0545 for FTS, CT11, and GG06 respectively). As was observed in the case of the ATM contracts, the FTS model systematically outperforms in terms of level of MCPDC for both the EUR-USD and EUR-GBP options. It correctly predicts the direction of implied volatility change up to 57.6% of the time (as seen for the 10 Delta one month EUR-GBP contract). The results for the EUR-JPY pair are more mixed however, with the CT11 model outperforming versus the FTS for 10, 25, and 75 delta option contracts. Again this mirrors the results of Table 2 in which we saw that modelling the intertemporal relationship between implied volatility smiles of EUR-JPY poses more of a difficulty for the FTS model than modelling the dynamics of EUR-GBP or EUR-USD. When looking at the asymmetric loss functions we find results mirroring the ATM case; the tendency over the period is for all models to under-predict the observations. The results suggest that the GG06 model is the least biased, a phenomenon that is most prominent in the 10 delta contracts of the three currencies. We have successfully compared and contrasted the three models in terms of calculated forecasting performance metrics. The FTS systematically outperforms CT11 and GG06 for all currencies in terms of minimising squared and absolute prediction errors. Furthermore, FTS outperforms in terms of directional accuracy for both EUR-USD and EUR-GBP, however it demonstrates some under-performance in terms of MCPDC for the most volatile currency, EUR-JPY. We now seek to formally test if this outperformance is sample specific or if we can draw inferences regarding the entire population. In line with a plethora of forecasting literature we adopt the Diebold-Mariano test statistic to establish the statistical significance of our findings.

We specify the hypotheses as follows:

$$H_0 : \theta_{benchmark} - \theta_{FTS} \leq 0$$

---

[9]Similar to Table 3 one-day ahead forecast performance metric results of 10, 25, 75, and 90 delta contract for three, six, and nine month maturities are available upon request.

Table 4: Diebold-Mariano Test EUR-USD

	FTS V CT11		FTS V GG06		FTS V AR	
h=1	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-12.87	0.00	-13.30	0.00	-3.03	0.00
3 month	-45.09	0.00	-13.73	0.00	0.66	0.75
6 month	-48.17	0.00	-13.83	0.00	-1.11	0.13
9 month	-48.21	0.00	-13.69	0.00	3.84	1.00
h=2	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-12.27	0.00	-11.24	0.00	-2.35	0.01
3 month	-43.03	0.00	-12.41	0.00	-1.65	0.05
6 month	-47.35	0.00	-13.70	0.00	-0.09	0.46
9 month	-47.87	0.00	-14.84	0.00	-5.20	0.00
h=3	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-11.50	0.00	-12.23	0.00	-5.80	0.00
3 month	-42.09	0.00	-13.85	0.00	-0.09	0.46
6 month	-47.47	0.00	-14.28	0.00	0.43	0.66
9 month	-47.68	0.00	-15.56	0.00	-5.07	0.00
h=5	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-15.95	0.00	-10.96	0.00	2.79	1.00
3 month	-41.67	0.00	-13.05	0.00	1.56	0.94
6 month	-47.28	0.00	-13.95	0.00	0.58	0.72
9 month	-47.80	0.00	-14.66	0.00	-1.70	0.04
h=10	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-15.31	0.00	-12.22	0.00	-4.70	0.00
3 month	-42.14	0.00	-13.76	0.00	-2.04	0.02
6 month	-47.57	0.00	-13.63	0.00	-1.72	0.04
9 month	-47.94	0.00	-14.56	0.00	-2.87	0.00

The table gives the cross model comparison results of applying the Diebold-Mariano (1995) test of predictive ability for h=1, 2, 3, 5, and 10 day ahead forecast horizons under a 500 day out-of-sample window length over the December 2011 to November 2013 period. The data is first aggregated across the implied volatility smile delta values for the EUR-USD currency pair. The values are provided for the Functional Time Series (FTS), CT11 model, GG06 model, and the autoregressive model (AR). Contract maturities of 1, 3, 6, and 9 months are analysed separately.

Table 5: Diebold-Mariano Test EUR-GBP

	FTS V CT11		FTS V GG06		FTS V AR	
h=1	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-12.58	0.00	-14.00	0.00	-4.98	0.00
3 month	-45.57	0.00	-14.69	0.00	-1.37	0.09
6 month	-49.23	0.00	-14.99	0.00	-3.09	0.00
9 month	-49.27	0.00	-15.36	0.00	-2.69	0.00
	FTS V CT11		FTS V GG06		FTS V AR	
h=2	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-10.37	0.00	-13.25	0.00	-5.62	0.00
3 month	-42.93	0.00	-14.18	0.00	-1.44	0.08
6 month	-48.97	0.00	-15.09	0.00	-2.17	0.02
9 month	-49.09	0.00	-14.97	0.00	-3.56	0.00
	FTS V CT11		FTS V GG06		FTS V AR	
h=3	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-12.35	0.00	-12.37	0.00	-8.63	0.00
3 month	-42.51	0.00	-13.24	0.00	-2.3	0.01
6 month	-48.63	0.00	-14.37	0.00	0.1	0.54
9 month	-48.82	0.00	-14.53	0.00	-0.31	0.38
	FTS V CT11		FTS V GG06		FTS V AR	
h=5	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-19.7	0.00	-14.52	0.00	-0.58	0.28
3 month	-43.97	0.00	-14.95	0.00	0.89	0.81
6 month	-48.85	0.00	-14.89	0.00	1.3	0.90
9 month	-48.92	0.00	-14.83	0.00	-0.02	0.49
	FTS V CT11		FTS V GG06		FTS V AR	
h=10	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-17.73	0.00	-13.28	0.00	-3.14	0.00
3 month	-43.05	0.00	-13.3	0.00	-2.58	0.01
6 month	-48.76	0.00	-13.75	0.00	-2.35	0.01
9 month	-48.81	0.00	-13.65	0.00	-2.37	0.01

The table gives the cross model comparison results of applying the Diebold-Mariano (1995) test of predictive ability for h=1, 2, 3, 5, and 10 day ahead forecast horizons under a 500 day out-of-sample window length over the December 2011 to November 2013 period. The data is first aggregated across the implied volatility smile delta values for the EUR-GBP currency pair. The values are provided for the Functional Time Series (FTS), CT11 model, GG06 model, and the autoregressive model (AR). Contract maturities of 1, 3, 6, and 9 months are analysed separately.

Table 6: Diebold-Mariano Test EUR-JPY

	FTS V CT11		FTS V GG06		FTS V AR	
h=1	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-10.29	0.00	-13.86	0.00	-0.02	0.49
3 month	-45.1	0.00	-14.34	0.00	0.95	0.83
6 month	-47.97	0.00	-14.39	0.00	-0.88	0.19
9 month	-47.85	0.00	-14.38	0.00	-0.27	0.4
	FTS V CT11		FTS V GG06		FTS V AR	
h=2	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-11.28	0.00	-13.79	0.00	2.90	1.00
3 month	-43.44	0.00	-13.79	0.00	-0.28	0.39
6 month	-47.71	0.00	-13.88	0.00	0.00	0.50
9 month	-47.76	0.00	-13.81	0.00	-2.97	0.00
	FTS V CT11		FTS V GG06		FTS V AR	
h=3	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-12.86	0.00	-13.88	0.00	-1.33	0.09
3 month	-44.92	0.00	-14.35	0.00	-1.02	0.15
6 month	-47.84	0.00	-14.37	0.00	-2.95	0.00
9 month	-47.76	0.00	-14.03	0.00	-2.13	0.02
	FTS V CT11		FTS V GG06		FTS V AR	
h=5	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-13.45	0.00	-13.52	0.00	-2.87	0.00
3 month	-44.36	0.00	-13.74	0.00	-3.95	0.00
6 month	-47.78	0.00	-14.09	0.00	-4.39	0.00
9 month	-47.75	0.00	-14.18	0.00	-3.71	0.00
	FTS V CT11		FTS V GG06		FTS V AR	
h=10	test statistic	p-value	test statistic	p-value	test statistic	p-value
1 month	-13.5	0.00	-14.46	0.00	-5.08	0.00
3 month	-43.26	0.00	-14.44	0.00	-5.23	0.00
6 month	-47.48	0.00	-14.38	0.00	-6.15	0.00
9 month	-47.49	0.00	-14.21	0.00	-6.21	0.00

The table gives the cross model comparison results of applying the Diebold-Mariano (1995) test of predictive ability for $h=1$, 2, 3, 5, and 10 day ahead forecast horizons under a 500 day out-of-sample window length over the December 2011 to November 2013 period. The data is first aggregated across the implied volatility smile delta values for the EUR-JPY currency pair. The values are provided for the Functional Time Series (FTS), CT11 model, GG06 model, and the autoregressive model (AR). Contract maturities of 1, 3, 6, and 9 months are analysed separately.

$$H_1 : \theta_{benchmark} - \theta_{FTS} > 0$$

where θ_{FTS} is the forecast error for the FTS model, and $\theta_{benchmark}$ is the forecast error for a given comparative benchmark model. The results of the tests for each currency are given in Tables 4, 5, and 6. For table brevity the $\theta_{benchmark}$ and θ_{FTS} forecasts are calculated for implied volatility data aggregated across the implied volatility smile for each contract maturity. This approach is similar to Chalamandaris and Tsekrekos (2014) who report p-values for all moneyness levels but only specific term structure segments. After applying the Diebold-Mariano test procedure, the FTS model demonstrates truly significant outperformance versus CT11 and GG06 in predicting EUR-USD, EUR-GBP, and EUR-JPY implied volatility for all 10, 25, 50, 75, and 90 deltas, and all 1, 3, 6, and 9 month option contracts, under the 500 day out-of-sample window period.[10] This result provides statistically confirmation of the outperformance we observe when comparing out-of-sample forecasting performance measure above. To make the comparison with prior literature more complete forecasting horizons of greater than one-day ahead, $h=2$, $h=3$, $h=5$, and $h=10$ are adopted. For these horizons we can also establish, using the Diebold-Mariano test, that statistically significant FTS outperformance versus CT11 and GG06 exists. Chalamandaris and Tsekrekos (2014) find that none of their proposed implied volatility models outperform the autoregressive benchmark in forecasts of less than five days ahead. This result is mirrored in Konstantinidi et al. (2008) and Dunis et al. (2013). As a result, in order to provide a further robustness check the performance of the FTS model is formally tested versus a standard univariate autoregressive model of order equal to the forecasting horizon. The autoregressive benchmark is applied to individual contracts separately. The intuition is that despite not taking account of implied volatility shape, the literature has shown such univariate models to constitute hard to beat benchmarks. In comparison to other implied volatility surface models however, our FTS model demonstrates improved performance, significantly outperforming the AR benchmark at a 10% level in 38 of the 60 cases tested. The model fails to reject the null hypothesis of equal performance or FTS under-performance in the remaining 22 cases. Additional testing indicates that the AR benchmark in fact significantly out performs the predictions from our proposed FTS model in only three cases.

4.4 Trading Strategy

Building on the evidenced predictability of the FTS model, we consider next how to exploit

---

[10] Similar results are observed for 100, 200, and 1000 day out-of-sample windows.

this information in a way that highlights the economic value of adopting this forecasting approach. To this end, we implement a stylised options trading strategy experiment designed to exploit the volatility predictions of the FTS model and we benchmark the performance against the CT11, GG06 and AR models. Following Bernales and Guidolin (2014), we utilise straddle trading strategies for our analysis as this gives exposure to movements in volatility while protecting against movements in the underlying FX rates. We proceed as follows. For a given currency pair and given maturity, we use the day-ahead prediction of the ATM volatility change under the FTS model as a signal to either buy or sell an ATM-straddle position of corresponding maturity. If the forecast is for volatility to increase then we go long the ATM-straddle and if the forecast is for volatility to decrease then we go short the ATM-straddle. As we do this for each of the 1, 3, 6 and 9 month maturities, leading to a portfolio of straddle positions for which we record the net daily return. The market is assumed frictionless with no transactions costs. The option pricing model of Garman and Kohlhagen (1983) is used to convert implied volatility quotes to prices, using the appropriate Euribor, USD Libor, GBP Libor and JPY Libor rates as required. We replicate this for each of the CT11, GG06 and AR models. Consistent with our earlier analysis, we consider 1000, 500, 200 and 100 day out-of-sample periods to assess the trading performance.

Following Chalamandris and Tsekrekos (2014), t-test results are presented for the null hypothesis that the trading strategy profits are a random sample from a normal distribution with zero mean and unknown variance, against the alternative that the mean is greater than zero. Table 7 summarises the results. It can be seen (from the results labelled *Full Sample*) that in the case of EUR-JPY, the FTS based straddle trading strategy leads to statistically significant profitability for each of the out-of-sample periods considered. Profitability is also identified in the case of EUR-USD for the 100-day out-of-sample period. In contrast, among the benchmark models, only the EUR-JPY trading strategy based on CT11 predictions is deemed profitable with statistical significance and this is only in the case of the 500-day out-of-sample period. None of the models are effective in trading the volatility of the EUR-GBP FX rate.

While the results appear quite favourable on the performance of the FTS model relative to the benchmark models, we note in our implementation a small number of large (in absolute terms) positive and negative trading strategy returns, which correspond to instances where the net value of the long/short straddle portfolio is very low. The implication of this is that the resulting changes in volatility over the next days lead to quite distorted returns. Such outliers have a direct impact on our ability to assess the performance of the trading strategies. We therefore objectively trim the trading to remove these tail outliers. Specifically, we remove 1% of the returns observations for

each out-of-sample period, split equally between the left and right tails. We refer to these updated samples as the *1%-Trimmed Samples* and we re-assess the trading performance based on these samples. For the FTS model, an improvement in performance is noted with the model predictions leading to statistically significant profitability for *both* the EUR-USD and EUR-JPY across all out-of-sample periods considered. In contrast, and despite accounting for the potential bias of outlier return observations, only the EUR-JPY trading strategy based on CT11 predictions is again deemed profitable with statistical significance and this is again for the 500 day out-of-sample period only. Once more, none of the models are effective in trading the volatility of the EUR-GBP FX rate.

The trading strategy results therefore provide incremental support for the effectiveness of the FTS forecasting model. The discussion highlights the economic value of adopting this functional forecasting approach. It is important though to qualify the stylised nature of the trading strategy experiment and note that further research would be required to comprehensively assess the potential for non-frictionless real world trading applications, most relevantly in markets with liquid exchange-based options availability. It is hoped that our work will motive further investigation in this direction.

5 Conclusion

We propose a functional time series (FTS) framework to characterise and forecast FX option implied volatility. Combining facets of models popular in implied volatility literature, namely Stock and Watson (2002) and the Gonçalves and Guidolin (2006) two-step forecasting framework we demonstrate a novel approach to forecasting the curve. Assuming continuous smoothness between adjacent moneyness points along the implied volatility smile differentiates the functional approach from a traditional discrete multivariate analysis. Parsimony is achieved through the use of functional principal component regression (FPCR), producing latent statistical factors that efficiently characterise the implied volatility process. Fitting the FTS model to our 2006-2013 data set produces a realistic and plausible implied volatility surface shape that is consistent with general equilibrium model theory. In comparison with the leading parametric models in the literature, Gonçalves and Guidolin (2006) (GG06) and Chalamandaris and Tsekrekos (2011) (CT11), the FTS model provides a superior fit across both the moneyness and term structure dimension. A major contribution of the study is that of demonstrating the performance advantage of adopting the FTS approach to predict future implied volatility movements. The performance of the proposed FTS based model

Table 7: Straddle Trading Strategy Testing

	1000-day out-of-sample period					
	Full Sample			*1%-Trimmed Sample*		
	USD	*GBP*	*JPY*	*USD*	*GBP*	*JPY*
FTS	1.04	1.05	1.94**	2.71***	1.06	2.34***
CT11	0.76	-1.72	0.70	0.29	-2.06	0.16
GG06	0.13	-0.68	1.10	-0.17	-0.87	0.73
AR	-1.49	0.19	0.39	-1.95	-3.22	0.98
	500-day out-of-sample period					
	Full Sample			*1%-Trimmed Sample*		
	USD	*GBP*	*JPY*	*USD*	*GBP*	*JPY*
FTS	1.03	1.03	2.43***	2.37***	0.47	2.48***
CT11	0.08	-0.06	2.23**	0.23	-0.24	2.44***
GG06	-0.09	-0.53	1.03	-0.03	-0.09	0.96
AR	-1.32	1.05	-0.31	-1.18	0.64	-0.65
	200 day out-of-sample period					
	Full Sample			*1%-Trimmed Sample*		
	USD	*GBP*	*JPY*	*USD*	*GBP*	*JPY*
FTS	1.02	1.00	2.41***	2.01**	0.49	2.42***
CT11	-1.46	-1.11	1.18	-1.64	-1.57	0.98
GG06	-0.33	-1.09	1.04	-0.32	-0.97	0.79
AR	-0.36	1.01	-0.71	-0.31	1.06	-1.21
	100 day out-of-sample period					
	Full Sample			*1%-Trimmed Sample*		
	USD	*GBP*	*JPY*	*USD*	*GBP*	*JPY*
FTS	1.75**	1.01	1.47*	1.75**	1.01	1.47*
CT11	-1.51	0.12	0.29	-1.51	0.12	0.29
GG06	-0.30	-0.41	-0.43	-0.30	-0.41	-0.43
AR	-0.62	1.01	-0.65	-0.62	1.01	-0.65

The table presents the results of the ATM-straddle trading strategies implemented in line with Bernales and Guidolin (2014) and described in Section 4.4. Following Chalamandris and Tsekrekos (2014), *t*-test results are reported for the null hypothesis that the trading strategy profits are a random sample from a normal distribution with zero mean and unknown variance, against the alternative that the mean is greater than zero. *Full Sample* refers to the full sample of return observations in the out-of-sample period. *1%-Trimmed Sample* refers to the sample of returns observations that result from objectively trimming 1% of the returns observations from the full sample, split equally between the left and right tails. The purpose of the trimming is to remove some noted extreme tail outliers that may effect our assessment of the trading strategy performance. Further discussion on this is given in Section 4.4.

is again benchmarked against the popular GG06 and CT11 models, with RMSE, MAE, MCPDC, MME(U), and MME(O) measures adopted. The Diebold-Mariano (1995) cross model predictability test is applied to validate the statistical significance of the FTS outperformance observed. To increase the robustness of the results various currencies, strikes, maturities, forecasting horizons, and out-of-sample windows are used.

References

Antoch, J., L. Prchal, M. R. De Rosa, and P. Sarda. 2008. *Functional linear regression with functional response: application to prediction of electricity consumption.* Springer.

Bekiros, S. D., and D. A. Georgoutsos. 2008. Non-linear dynamics in financial asset returns: the predictive power of the CBOE volatility index. *The European Journal of Finance* 14:397–408.

Benko, M., W. Härdle, and A. Kneip. 2009. Common functional principal components. *The Annals of Statistics* 37:1–34.

Bernales, A., and M. Guidolin. 2014. Can we forecast the implied volatility surface dynamics of equity options? Predictability and economic value tests. *Journal of Banking & Finance* 46:326–342.

Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *The Journal of Political Economy* pp. 637–654.

Brailsford, T. J., and R. W. Faff. 1996. An evaluation of volatility forecasting techniques. *Journal of Banking & Finance* 20:419–438.

Chalamandaris, G., and A. E. Tsekrekos. 2010. Predictable dynamics in implied volatility surfaces from OTC currency options. *Journal of Banking & Finance* 34:1175–1188.

Chalamandaris, G., and A. 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, G., and A. E. Tsekrekos. 2014. Predictability in implied volatility surfaces: evidence from the Euro OTC FX market. *The European Journal of Finance* 20:33–58.

Corrado, C. J., and T. W. Miller. 2006. Estimating expected excess returns using historical and option-implied volatility. *Journal of Financial Research* 29:95–112.

David, A., and P. Veronesi. 2000. Option prices with uncertain fundamenftals: Theory and evidence on the dynamics of implied volatilities .

Diebold, F. X., and C. Li. 2006. Forecasting the term structure of government bond yields. *Journal of econometrics* 130:337–364.

Diebold, F. X., and R. S. Mariano. 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* 13:253–263.

Dumas, B., J. Fleming, and R. E. Whaley. 1998. Implied volatility functions: Empirical tests. *The Journal of Finance* 53:2059–2106.

Fuertes, A.-M., M. Izzeldin, and E. Kalotychou. 2009. On forecasting daily stock volatility: the role of intraday information and market conditions. *International Journal of Forecasting* 25:259–281.

Garcia, R., R. Luger, and E. Renault. 2003. Empirical assessment of an intertemporal option pricing model with latent variables. *Journal of Econometrics* 116:49–83.

Garman, M. B., and S. W. Kohlhagen. 1983. Foreign currency option values. *Journal of International Money and Finance* 2:231–237.

Garvey, J. F., and L. A. Gallagher. 2012. The Realised–Implied Volatility Relationship: Recent Empirical Evidence from FTSE-100 Stocks. *Journal of Forecasting* 31:639–660.

Goncalves, S., and M. Guidolin. 2006. Predictable dynamics in the S&P 500 index options implied volatility surface*. *The Journal of Business* 79:1591–1635.

Guidolin, M., and A. Timmermann. 2003. Option prices under Bayesian learning: implied volatility dynamics and predictive densities. *Journal of Economic Dynamics and Control* 27:717–769.

Horváth, L., and P. Kokoszka. 2012. *Inference for functional data with applications*, vol. 200. Springer.

Hyndman, R. J., and H. L. Shang. 2009. Forecasting functional time series. *Journal of the Korean Statistical Society* 38:199–211.

Kargin, V., and A. Onatski. 2008. Curve forecasting by functional autoregression. *Journal of Multivariate Analysis* 99:2508–2526.

Kearney, F., F. Murphy, and M. Cummins. 2015. An analysis of implied volatility jump dynamics: Novel functional data representation in crude oil markets. *North American Journal of Economics and Finance* 33:199–216.

Konstantinidi, E., G. Skiadopoulos, and E. Tzagkaraki. 2008. Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices. *Journal of Banking & Finance* 32:2401–2411.

Kosiorowski, D. 2014. Functional Regression in Short-Term Prediction of Economic Time Series. *Statistics in Transition new series* 15:611–626.

Liebl, D. 2013. Modeling and forecasting electricity spot prices: A functional data perspective. *The Annals of Applied Statistics* 7:1562–1592.

Müller, H.-G., R. Sen, and U. Stadtmüller. 2011. Functional data analysis for volatility. *Journal of Econometrics* 165:233–245.

Muzzioli, S. 2010. Option-based forecasts of volatility: an empirical study in the DAX-index options market. *The European Journal of Finance* 16:561–586.

Ramsay, J., and B. Silverman. 2005. *Functional data analysis.* Springer, New York.

Shang, H. L. 2013. Functional time series approach for forecasting very short-term electricity demand. *Journal of Applied Statistics* 40:152–168.

Shang, H. L. 2015. Forecasting Intraday S&P 500 Index Returns: A Functional Time Series Approach. URL `http://ssrn.com/abstract=2647233`.

Stock, J. H., and M. W. Watson. 2002. Forecasting using principal components from a large number of predictors. *Journal of the American statistical association* 97:1167–1179.

Taylor, S. J., P. K. Yadav, and Y. Zhang. 2010. The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks. *Journal of Banking & Finance* 34:871–881.

Yu, W. W., E. C. Lui, and J. W. Wang. 2010. The predictive power of the implied volatility of options traded OTC and on exchanges. *Journal of Banking & Finance* 34:1–11.