Forecasting global stock market implied volatility indices

Degiannakis, Stavros and Filis, George and Hassani, Hossein

Panteion University of Social and Political Sciences, Bournemouth University, Institute for International Energy Studies

1 September 2015

Forecasting global stock market implied volatility indices

Stavros Degiannakis[1,2], George Filis[3]*, Hossein Hassani[4]

[1]Department of Economics and Regional Development, Panteion University of Social and Political Sciences, 136 Syggrou Avenue, 17671, Greece.

[2]Postgraduate Department of Business Administration, Hellenic Open University, Aristotelous 18, 26 335, Greece.

[3]Bournemouth University, Department of Accounting, Finance and Economics, Executive Business Centre, 89 Holdenhurst Road, BH8 8EB, Bournemouth, UK.

[4]Research Institute of Energy Management and Planning, University of Tehran, No. 13, Ghods St., Enghelab Ave., Tehran, Iran.

*Corresponding author: email: gfilis@bournemouth.ac.uk, tel: 0044 (0) 01202968739, fax: 0044 (0) 01202968833

Abstract

This study compares parametric and non-parametric techniques in terms of their forecasting power on implied volatility indices. We extend our comparisons using combined and model-averaging models. The forecasting models are applied on eight implied volatility indices of the most important stock market indices. We provide evidence that the non-parametric models of Singular Spectrum Analysis combined with Holt-Winters (SSA-HW) exhibit statistically superior predictive ability for the one and ten trading days ahead forecasting horizon. By contrast, the model-averaged forecasts based on both parametric (Autoregressive Integrated model) and non-parametric models (SSA-HW) are able to provide improved forecasts, particularly for the ten trading days ahead forecasting horizon. For robustness purposes, we build two trading strategies based on the aforementioned forecasts, which further confirm that the SSA-HW and the ARI-SSA-HW are able to generate significantly higher net daily returns in the out-of-sample period.

Keywords: Stock market, Implied Volatility, Volatility Forecasting, Singular Spectrum Analysis, ARFIMA, HAR, Holt-Winters, Model Confidence Set, Model-Averaged Forecasts.

JEL codes: C14; C22; C52; C53; G15.

1. Introduction and review of the literature

It has been well established that stock market volatility forecasting is important for investors, portfolio managers, asset valuation, hedging strategies, risk management purposes, as well as, policy makers (see, *inter alia*, Figlewski, 1997; Andersen *et al.*, 2003,2005; Christodoulakis, 2007; Fuertes *et al.*, 2009; Charles, 2010; Barunik *et al.*, 2016).

For instance, investors and portfolio managers seek a prediction of their future uncertainty in order to estimate a specific upper limit of risk that are willing to accept, to reach optimal portfolio decisions and to form appropriate hedging strategies.

Even more, forecasting volatility is the single most important component for pricing derivative products, such as option contracts. Unless derivatives contracts are priced correctly, hedging strategies can be expensive and not yield the desired outcome. Nowadays, volatility can be the underlying asset of derivatives products, such as in the VIX futures contracts. Thus, forecasting the expected volatility of the underlying asset helps for the correct valuation of these contracts.

Forecasting volatility is also important for policy makers, since it informs monetary policy decisions and it allows for measuring the expectations of the financial markets regarding the (un)successful outcome of fiscal and/or monetary policy decisions. The aforementioned arguments render important the accurate stock market volatility forecasting.

The vast majority of the stock market volatility forecasting studies have concentrated their attention on the use of models which are variants of GARCH models (see, *inter alia*, Bollerslev *et al.*, 1994; Degiannakis, 2004; Hansen and Lunde, 2005), stochastic volatility models (see, among others, Deo, 2006; Yu, 2012) or realized volatility models (Andersen *et al.*, 2003, Andersen *et al.*, 2005).

These models generate forecasts of the current looking volatility, despite the fact that implied volatility indices have been long considered as better predictors of the future volatility (see for instance, Chiras and Manaster, 1978; Beckers, 1981). More recently, studies by Fleming *et al.* (1995), Christensen and Prabhala (1998), Fleming (1998), Blair *et al.* (2001), Simon (2003), Giot (2003), Degiannakis (2008a) and Frijns *et al.* (2008a) have also provided evidence that implied volatility is more informative when we forecast stock market volatility.

Methodologically, the literature provides evidence that the fractionally integrated autoregressive moving average models outperform the volatility forecasts that are produced by the GARCH and stochastic volatility models (Koopman *et al.*, 2005). Degiannakis (2008b) also maintains that due to the long memory property of volatility, the ARFIMA framework is suitable for estimating and forecasting the logarithmic transformation of volatility. At the same time, some argue that heterogeneous autoregressive models (HAR) are more successful in forecasting volatility due to the fact that they are parsimonious and they can capture the long-memory that is observed in volatility (see, *inter alia*, Andersen *et al.*, 2007; Corsi, 2009; Busch *et al.*, 2011; Fernandes *et al.*, 2014, Sevi, 2014). Nevertheless, Angelidis and Degiannakis (2008) provide evidence that there is not a unique model that is offering better predictive ability than others in all instances.

Despite the fact that the existing evidence has established that models such as ARFIMA and HAR are the best performing forecasting models, the literature remains relatively silent in the use of various non-parametric techniques when forecasting stock market implied volatility.

The rather limited literature on volatility forecasting using non-parametric techniques or a combination of parametric and non-parametric techniques provides some encouraging results, although it concentrates its attention on the use of biological algorithms and neural networks. For instance, Hung (2011a,b) combines fuzzy systems with the GARCH models and shows that such combinations provide significant predictive gains. Wei (2013) provides similar findings using an adaptive network-based fuzzy inference system (ANFIS), employing genetic algorithms to calibrate the weights of the rules in the ANFIS model. Furthermore, several authors combine artificial neural networks (ANN) with GARCH-type models to forecast stock market volatility and their findings corroborate the ones presented before, suggesting that such combinations could lead to significant reduction in the predictive error of parametric models (see, *inter alia*, Kristjanpoller *et al.*, 2014; Hajizadeh *et al.*, 2012; Bildirici and Ersin, 2009, Donaldsona and Kamstrab, 1997).

Adding to this literature we focus on the use of Singular Spectrum Analysis (SSA) in forecasting stock market volatility. SSA is regarded as a non-parametric technique for time series analysis and forecasting, which offers great success in forecasting economic and financial series (see for example, Hassani *et al.*, 2009;

Beneki *et al.*, 2012). Nevertheless, it has not been applied before to the forecast of implied volatility indices, despite the fact that since the early 2000s Thomakos *et al.* (2002) maintained that SSA is able to decompose volatility series more effectively, capturing both the market trend and a number of market periodicities. Thus, an important extension to the existing literature would be to assess the forecasting ability of SSA in the context of volatility modeling.

Overall, the limited empirical applications of SSA to economic and financial series provide so far significant evidence of its superior predictive ability against the standard forecasting models, such as the ARIMA-type and GARCH-type models.

In short, SSA decomposes a time series into the sum of a small number of independent and interpretable components such as a slowly varying trend, oscillatory components and noise (Hassani *et al.*, 2009). The main advantage of SSA-type models is that they do not require any statistical assumptions in terms of the stationarity of the series or the distribution of the residuals. In fact, SSA uses bootstrapping to generate the confidence intervals that are required for the evaluation of the forecasts (Hassani and Zhigljavsky, 2009; Vautard *et al.*, 1992).

The aim of this study is to use both the best parametric forecasting techniques (such as ARFIMA and HAR) and the best performing non-parametric forecasting techniques (such as SSA) in the forecast of implied volatility indices. We further our comparisons using model-averaging forecasts. For robustness purposes, we compare the forecasts from the aforementioned models with four naïve models; i.e. I(1), ARI(1,1), FI(1) and ARFI(1,1). The forecasting horizons are 1-day and 10-days ahead and they are chosen as these time horizons are more adequate for investors and portfolio managers, according to the aforementioned volatility forecasting literature.

The contribution of the paper is described succinctly. First, we provide an alternative model to forecast implied volatility; second, we open new avenues for the use of SSA-type in finance and third, we contribute to the non-parametric literature of financial markets.

The study provides empirically significant evidence that the combination of two non-parametric models (SSA and Holt-Winter (HW)) achieves more accurate forecasts for the 1-day and 10-days ahead, compared to the parametric models of ARFIMA, HAR, as well as, to the four naïve models. On the other hand, model-averaged forecasts reveal that the forecasting accuracy of the SSA-HW is enhanced, particularly for the 10-days ahead, if it is combined with the ARI(1,1) model. The predictive accuracy is assessed by the Mean Squared Error (MSE) and the Mean Absolute Error (MAE) loss functions, the Model Confidence Set forecasting evaluation procedure and the Direction-of-Change criterion. Finally, we assess the forecasting ability of the models by means of two trading strategies. The results reveal that investors can generate significant positive average net profits using the SSA-HW and the ARI-SSA-HW models.

The rest of the paper is structured as follows. Section 2 presents the data of the study, followed by Section 3, which illustrates the forecasting framework. Section 4 provides a detailed explanation of the implied volatility forecasts estimation procedure and section 5 describes the adopted forecasting evaluation methods. Section 6 analyses the empirical findings, whereas Section 7 concludes the study.

2. Data description

We use daily data from the 1st of February, 2001 up to the 9th of July, 2013 (i.e. 3132 trading days) from eight implied volatility indices. The implied volatilities are the following: VIX (S&P500 Volatility Index – US), VXN (Nasdaq-100 Volatility Index – US), VXD (Dow Jones Volatility Index – US), VSTOXX (Euro Stoxx 50 Volatility Index – Europe), VFTSE (FTSE 100 Volatility Index – UK), VDAX (DAX 30 Volatility Index – Germany), VCAC (CAC 40 Volatility Index – France) and VXJ (Japanese Volatility Index - Japan). The stock markets under consideration represent six out of the ten most important stock markets internationally, in terms of capitalization. In addition, these markets are among the most liquid markets of the world. Thus, we maintain that their implied volatility indices are representative of the world's stock market uncertainty. The data were extracted from *Datastream*®. As we aim for a common sample of the aforementioned implied volatility indices, the starting data of the sample period were dictated by the availability of the data of the VXN index.

Figure 1 and Table 1 exhibit the series under consideration and list their descriptive statistics, respectively.

[FIGURE 1 HERE]
[TABLE 1 HERE]

In Figure 1 we observe that all implied volatility indices display very similar patterns. For example, it is evident that during the Great Recession of 2007-2009 all indices reached their highest level over the sample period. In addition, the magnitude of these peaks is comparable across indices. Furthermore, we observe two more peaks in 2003 and 2011, respectively. The volatility spikes in 2003 can be attributed to the second war in Iraq, whereas a plausible explanation of the 2011 peak in stock market volatilities can be found in the European debt crisis which initiated in Greece before spreading to other countries such as Ireland, Spain and Portugal. The US debt-ceiling crisis of the same year could have aggravated higher uncertainty in world stock markets.

In Table 1 we notice that average volatility is of similar size across indices, with the exception being the VXN and VXD indices, which exhibit the highest and lowest average volatility, respectively. Furthermore, the VXN index also exhibits the highest level of standard deviation, suggesting that it is the most volatile index. All series under examination are stationary and heteroscedastic, as suggested by the ADF and ARCH LM tests, respectively.

3. Methodology and IV-SSA-HW model

The modelling and forecasting of economic and financial time series are often rendered difficult due to their non-stationary nature and frequent structural breaks. In this light, the SSA technique can be particularly advantageous as it is not bound by the assumptions of stationarity, linearity and normality, which govern classical time series analysis and forecasting models (Hassani *et al.*, 2017). As a result, we can obtain a comparatively more realistic approximation to the real data. Moreover, unlike classical models, which forecast both the signal and noise in tandem, the SSA has the capacity to extract a more accurate signal from the implied volatility series and thus helps to improve the accuracy of the final forecast (Hassani and Thomakos, 2010). Furthermore, unlike parametric forecasting models which rely on several unknown parameters, the SSA technique relies solely on the choices of its Window Length, L and the number of eigenvalues, r. The SSA technique has also proven to be a viable option for forecasting during recessions, when faced with structural breaks in time series (see for example, Hassani *et al.,* 2013; Silva and Hassani, 2015). Relevant to the aforementioned point, it is also worth noting that SSA can handle both short and long time series equally successfully where classical methods fail (Silva and Hassani, 2015).

Obviously, there are several linear and nonlinear filtering methods such as the Hodrick-Prescott filter, ARMA model, simple nonlinear filtering and local projective. However, the SSA technique relies on the Singular Value Decomposition (SVD) approach for noise reduction, which is regarded as a more effective noise reduction tool in comparison to standard filtering techniques which decompose series in different frequencies (Soofi and Cao, 2002; Ortu *et al.*, 2013). Furthermore, unlike local methods, such as linear filtering or wavelets, or even the HW, the SSA exploits the trajectory matrix computed using all parts of a time series (Alexandrov, 2009). In the past, one of the main drawbacks of the SVD approach was its computational complexity. However, the use of modern day technology and parallel algorithms have helped to reduce this shortcoming (Golyandina *et al.,* 2015).

In this paper, we combine the advantages of SSA as a filtering method, along with Holt Winters' (HW) non-parametric forecasting capacity. Whilst it is possible to build a combination forecast using any other time series analysis and forecasting technique, here we opted for SSA in combination with HW as HW, similar to SSA, is a non-parametric technique. Accordingly, by combining two non-parametric techniques, we can clear out the need for assumptions that must be considered when adopting parametric techniques.

To motivate further the combination of SSA-HW, we turn our attention to the stylized facts of volatility. For instance, (i) implied volatility indices are highly persistent, (ii) the autocorrelations of the index level and the logarithm of the index level are statistically significant and positive for at least 250 trading days and (iii) implied volatility indices are mean reverting in the long run. Thus, changes in volatility have a very long-lasting impact on its subsequent evolution. ARFIMA and HAR models are trying to capture that type of long memory property. However, the SSA can decompose the implied volatility series more effectively, capturing both the market trend and the volatility periodicities.

In addition, volatility is not constant and tends to cluster through time. Observing a large (small) implied volatility today is a good precursor of large (small) implied volatility in the coming days. HW is an appropriate forecasting technique for series with a time trend and additive (or multiplicative) periodic variation. The HW technique is characterised by its ability to decompose non-parametrically the forecasting procedure into the smoothing equation for the level of the predicted series, the trend equation and the periodic component.

Furthermore, the SSA-HW combination allows a compromise between model parsimony and forecast accuracy. In brief, the principle of parsimony suggests that one must opt for the model with the smallest number of parameters (simplest model) such that an adequate representation of the actual data is provided (Chatfield, 1996). When combining forecasts, studies indicate that forecasting accuracy can only be improved if forecasts are combined from two adequate parsimonious forecasting models (McLeod, 1993). Parsimony also allows better predictions and generalizations of new data as it helps to distinguish the signal from the noise (Busemeyer *et al.*, 2015). This is in addition to the preference for parsimony as an approach for avoiding over-parameterization when modelling data for forecasting (Booth and Tickle, 2008) and it is a recommended criterion for differentiating between forecasting models (Harvey, 1990). However, the best compromise between model parsimony and forecast accuracy is likely to consider whether the forecasts from the parsimonious model are significantly more accurate than a forecast from a competing model, provided the models in question are not affected by over or under fitting.

Thus, in this paper, even though we decompose the implied volatility series using SSA and we then forecast each of the decomposed series using the HW model[1], we also forecast each of the implied volatility series using the SSA and HW separately.

In the decomposition stage, the first step is referred to the embedding process and the construction of the trajectory matrix. Consider the implied volatility index IV_t of length $\breve{T}$. Embedding process maps the one dimensional time series IV_t into a multidimensional time series $X_1,..., X_K$ with vectors $X_i = [IV_i, IV_{i+1}, IV_{i+2},..., IV_{i+\mathbf{L}-1}]'$, where $\mathbf{L}$ is an integer such that $2 \leq \mathbf{L} \leq \breve{T} - 1$. The selection of the optimal window length $\mathbf{L}$ for decomposing the time series is based on the RMSE criterion[2]. The

[1] The SSA-HW model is estimated in R software.

[2] The implied volatility series is divided into training and test sets. Decomposition of the training set is evaluated for different window lengths and eigenvalues. The results from the best decomposition as determined via the training approach is then used to decompose the test set of each index and then forecasted individually with HW prior to combining these decomposed forecasts for which the out-of-sample forecasting errors are reported.

trajectory matrix, $\mathbf{X}$, is constructed such that $K = \breve{T} - \mathbf{L} + 1$; $\mathbf{X}$ is a Hankel matrix, i.e. elements along the diagonal $i+j$ equal:

$$\mathbf{X} = [X_1, ..., X_r, ..., X_K] = (x_{i,j})_{i,j=1}^{\mathbf{L},K} = \begin{pmatrix} IV_1 & IV_2 & IV_3 & \cdots & IV_K \\ IV_2 & IV_3 & IV_4 & \cdots & IV_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ IV_{\mathbf{L}} & IV_{\mathbf{L}+1} & IV_{\mathbf{L}+2} & \cdots & IV_{\breve{T}} \end{pmatrix}. \quad (1)$$

The second step of the decomposition stage is known as singular value decomposition (SVD). In order to obtain the SVD of the trajectory matrix $\mathbf{X}$, we calculate $\mathbf{XX}'$ for which $\lambda_1, ..., \lambda_{\mathbf{L}}$ denote the eigenvalues in decreasing order, and $U_1, ..., U_{\mathbf{L}}$ represent the corresponding eigenvectors. The SVD step then provides the singular values r (the second parameter of SSA), such that $\mathbf{X} = X_1 + ... + X_r$. Thereafter, we use diagonal averaging to transform the components of the matrix $\mathbf{X}$ into a Hankel matrix which can then be converted into time series $IV_{t,1} \ldots . IV_{t,r}$, where $IV_{t,r}$ refers to the decomposed time series from the original implied volatility index. Having decomposed the implied volatility series, we apply the HW algorithm (Hyndman *et al.*, 2013) to forecast the decomposed series $IV_{t,1} \ldots . IV_{t,r}$.

In this paper, during the SSA filtering process, we follow a binary approach and extract the trend and two other leading components (henceforth, r=3) whilst considering the remaining components as noise, in line to the standard practice in SSA applications (Hassani *et al.*, 2017)[3].

We propose the combination of the forecasts attained via HW for each decomposed component via aggregation. The underlying idea behind this approach is to firstly decompose a given series, so that we can identify the various fluctuations, which were previously hidden under the overall series and secondly, to forecast each of these decompositions with HW. In this way, the model can capture all fluctuations, which were hidden previously, and then combine all these forecasts via aggregation to generate the SSA-HW forecast. Depending on the characteristics of the time series, the Hyndman *et al.* (2013) algorithm automatically selects either the multiplicative or

[3] The extracted components are available upon request.

the additive HW method. The additive HW framework for forecasting the decomposed series, $IV_{t,r}$, is presented as:

$$\begin{aligned}
\hat{l}_{t,r} &= \hat{\alpha}_r\left(IV_{t,r} - \hat{s}_{t-m,r}\right) + (1-\hat{\alpha}_r)\left(\hat{l}_{t-1,r} + \hat{b}_{t-1,r}\right) \\
\hat{b}_{t,r} &= \hat{\beta}_r\left(\hat{l}_{t,r} - \hat{l}_{t-1,r}\right) + \left(1-\hat{\beta}_r\right)\hat{b}_{t-1,r} \\
\hat{s}_{t,r} &= \hat{\gamma}_r\left(IV_{t,r} - \hat{l}_{t-1,r} - \hat{b}_{t-1,r}\right) + (1-\hat{\gamma}_r)\hat{s}_{t-m,r},
\end{aligned} \tag{2}$$

where $\hat{l}_{t,r}$ is the smoothing equation for the level, $b_{t,r}$ is for the trend, $s_{t,r}$ is the periodicity equation and m is used to denote the periodicity frequency. The alternative, which is the multiplicative HW method has the form:

$$\begin{aligned}
\hat{l}_{t,r} &= \hat{\alpha}\left(IV_{t,r} / \hat{s}_{t-m,r}\right) + (1-\hat{\alpha})\left(\hat{l}_{t-1,r} + \hat{b}_{t-1,r}\right) \\
\hat{b}_{t,r} &= \hat{\beta}_r\left(\hat{l}_{t,r} - \hat{l}_{t-1,r}\right) + \left(1-\hat{\beta}_r\right)\hat{b}_{t-1,r} \\
\hat{s}_{t,r} &= \hat{\gamma}_r\left(IV_{t,r} / \hat{l}_{t,r}\right) + (1-\hat{\gamma}_r)\hat{s}_{t-m,r}.
\end{aligned} \tag{3}$$

4. Forecasting IV indices

4.1. IV-SSA-HW model

We aggregate the Holt-Winters forecasts obtained for time series $IV_{t,1}$…. $IV_{t,r}$ to arrive at the SSA-HW forecasts. The additive HW one-step-ahead, $IV_{t+1|t}$, and 10-days-ahead, $IV_{t+10|t}$, implied volatility forecasts are computed as:

$$IV_{t+1|t} = \sum_{r=1}^{3}\left(\hat{l}_{t,r} + \hat{b}_{t,r} + \hat{s}_{t+1-m,r}\right) \tag{4}$$

and

$$IV_{t+10|t} = \sum_{r=1}^{3}\left(\hat{l}_{t,r} + 10\,\hat{b}_{t,r} + \hat{s}_{t+10-m,r}\right), \tag{5}$$

respectively. By contrast, the multiplicative HW one-step-ahead, $IV_{t+1|t}$, and 10-days-ahead, $IV_{t+10|t}$, implied volatility forecasts are computed as:

$$IV_{t+1|t} = (\hat{l}_{t,r} + \hat{b}_{t,r}) * \hat{s}_{t+1-m,r} \tag{6}$$

and

$$IV_{t+10|t} = (\hat{l}_{t,r} + 10\,\hat{b}_{t,r}) * \hat{s}_{t+10-m,r}, \tag{7}$$

respectively[4].

4.2. Naïve models, ARFIMA, HAR & model-averaged forecasts

As mentioned in Section 1, apart from the model frameworks presented in this section we further employ four naïve models, namely, the I(1), ARI(1,1), FI(1) and ARFI(1,1), the HW and SSA models, separately, as well as, the ARFIMA and HAR models. For brevity, these models' specifications are presented in the Appendix.

Furthermore, we employ model-averaged forecasts combining the best naïve model with the HAR, ARFIMA and SSA-HW. In addition, since the aim of the study is to compare non-parametric models and their combination against parametric models, we also proceed with the model-averaged forecast of the HAR-ARFIMA model. Forecasting literature states (i.e. Favero and Aiolfi, 2005, Samuels and Sekkel, 2013, Timmermann, 2006) that model-averaged forecasts provide incremental predictive gains compared to single models. In particular, forecast combinations with (i) equal weight averaging and (ii) fewer models included in the combination provide more accurate forecasts.

Even though the literature suggests that equal weight averaging may work particularly well, we also consider the Granger and Ramanathan (1984) approach, where the weights of the model average forecasts are based on their forecasting performance in the most recent past. The combined forecasts $IV_{t+s|t,(c)}$ are computed recursively as follows:

$$IV_{t+s|t,(c)} = w_{0,(t)} + w_{1,(t)} IV_{t+s|t,(1)} + w_{2,(t)} IV_{t+s|t,(2)}, \tag{8}$$

where $IV_{t+s|t,(1)}$ and $IV_{t+s|t,(2)}$ are the s-step-ahead forecasts from models (1) and (2), whereas the $w_{0,(t)}$, $w_{1,(t)}$ and $w_{2,(t)}$ denote the OLS recursive estimates from $IV_t = w_{0,(t)} + w_{1,(t)} IV_{t|t-s,(1)} + w_{2,(t)} IV_{t|t-s,(2)} + u_t$, for $u_t \sim N(0, \sigma_u^2)$.

In order to avoid a forward looking bias, at each trading day *t*, the weights are re-estimated based on the 250 most recent past forecasts. The intercept $w_{0,(t)}$ allows for a possible bias adjustment in the combined forecast. The combined forecasts have been also computed (i) without the intercept and (ii) for the sum of weights to equal 1 (i.e. $w_{1,(t)} + w_{2,(t)}$ =1). Nevertheless, the latter two approaches, and the equally

[4] For the calibration and estimation of the HW parameters, please see Hyndman and Athanasopoulos (2014).

weighted combined forecasts did not achieve better forecasts (which is in line with Granger and Ramanathan, 1984), thus, we only present the combined forecasts based on Eq 8.

5. Forecasting evaluation

5.1. MSE, MAPE loss functions and the model confidence set

The training period of the models is $\tilde{T}$ =1000 days, i.e. from 02/02/2001 until 28/01/2005[5]. The remaining T =2132 days are used for the evaluation period of the out-of-sample forecasts. In order to proceed to the first out-of-sample forecast (i.e. *t+1* forecast or day 1001), we train the models using the initial 1000 days. A rolling window approach with fixed length of 1000 days is used for all subsequent forecasts. The use of a restricted window length of 1000 trading days incorporates changes in trading behaviour more efficiently. For example, Angelidis *et al.* (2004), Degiannakis *et al.* (2008) and Engle *et al.* (1993) provide empirical evidence that the use of restricted rolling window samples captures the changes in market activity more effectively[6,7]. The total number of observations is $\breve{T} = \tilde{T} + T$. The forecasting accuracy of the models is initially gauged using two established loss functions, the Mean Squared Error, $MSE = T^{-1}\sum_{t=1}^{T}\left(IV_{t+n|t} - IV_{t+n}\right)^2$, and the Mean Absolute Error, $MAE = T^{-1}\sum_{t=1}^{T}\left|IV_{t+n|t} - IV_{t+n}\right|$, where, $IV_{t+n|t}$ is the implied volatility forecast, whereas IV_{t+n} is the actual implied volatility .[8]

[5] There are two reasons that justify the choice of initial training period. First, a large sample size for the estimation of the models was required. Second, it was preferable for our initial training period to stop before the Global Financial Crisis of 2007-09. The inclusion of the Global Financial Crisis period in the out-of-sample period allows for the better evaluation of the forecasting models' performance. Nevertheless, a training period of 750 and 1250 days was also considered and the results are qualitatively similar.

[6] For robustness, we used various window lengths for the rolling window approach and the results remain qualitatively unchanged.

[7] We also considered a recursive approach, where for each subsequent forecast after the $t+1$ forecast we added an additional day to the training period. For example, for the $t+2$ forecast we used $\tilde{T}+1$ daily observations. The results are qualitatively similar and they are available upon request.

[8] An alternative forecasting evaluation method is the Mincer and Zarnowitz (1969) regression, where the future VIX is regressed against the three different forecasts. The coefficients of the regressions are interpreted as the amount of information embedded in the different forecasts. The results are qualitatively similar.

In addition, we employ the Model Confidence Set (MCS) procedure of Hansen *et al.* (2011). The MCS test determines the set of models that consists of the best models where best is defined in terms of a predefined loss function. In our case two loss functions are employed, namely the MSE and the MAE. The MCS compares the predictive accuracy of an initial set of M^0 models and investigates, at a predefined level of significance, which models survive the elimination algorithm. For $L_{i,t}$ denoting the loss function of model i at day t, and $d_{i,j,t} \equiv L_{i,t} - L_{j,t}$ is the evaluation differential for $i, j \in M^0$ the hypotheses that are being tested are:

$$H_{0,M} : E\left(d_{i,j,t}\right) = 0 \tag{9}$$

for $\vee\, i, j \in M$, $M \subset M^0$ against the alternative $H_{1,M} : E\left(d_{i,j,t}\right) \neq 0$ for some $i, j \in M$. The elimination algorithm based on an equivalence test and an elimination rule, employs the equivalence test for investigating the $H_{0,M}$ for $\forall M \subset M^0$ and the elimination rule to identify the model i to be removed from M in the case that $H_{0,M}$ is rejected.

We should highlight here that several studies compare their forecasting models against a pre-selected benchmark, using tests, such as the Diebold-Mariano (Diebold and Mariano, 1995) for pairwise comparisons, the Equal Predictive Accuracy test (Clark and West, 2007) for nested models, or even the Reality Check for Data Snooping (White, 2000) and the Superior Predictive Ability (Hansen, 2005) for multiple comparisons.

By contrast, in this case we are not interested in pairwise comparisons, nor we have a benchmark model as the aim is to simultaneously evaluate the forecasting performance of the competing models and evaluate which models belong to the set of the best performing models.

In any case, the Superior Predictive Ability (SPA) test of Hansen (2005) was also used to evaluate the forecasting accuracy of the competing models, for robustness purposes. Initially, the benchmark model for the SPA test was the ARI(1,1), which is the best naïve model. Subsequently, we used the IV-HAR and the IV-ARFIMA as benchmark models against the SSA-HW. The results confirm the MCS findings and although they are not reported here, they are available upon request.

5.2. Direction-of-change

Furthermore, we consider the Direction-of-Change (DoC) forecasting evaluation technique. The DoC is particularly important for trading strategies as it provides an evaluation of the market timing ability of the forecasting models. The DoC criterion reports the proportion of trading days that a model correctly predicts the direction (up or down) of the volatility movement for the 1-day and 10-days ahead.

5.3. Forecast evaluations based on trading strategies

Finally, we compare the performance of each forecasting method based on two trading strategies. In the first trading strategy, the investor invests into a single-asset portfolio, which is composed by an implied-volatility index (i.e. we assume that each implied volatility index is a tradable asset). For the 1-day ahead forecasts, the trader takes a long position when the $t+1$ forecasted implied volatility of model i is higher compared to the actual implied volatility at time t. By contrast, when the $t+1$ forecasted implied volatility of model i is lower compared to the actual implied volatility at time t, then the trader takes a short position. Put it simply, when the investor expects an implied volatility index to increase (decrease) at $t+1$ based on model i then she goes long (short) in the specific implied volatility index. Similarly, we construct the trading strategy for the 10-days ahead forecasts. Portfolio returns are computed as the average net daily returns over the investment horizon, which coincides with our out-of-sample forecasting period of T =2132 days. The transaction costs per unit for each trade are estimated to be between 0.6%-1.2% (see Jung, 2016). The intuition of this rather naïve trading strategy is to evaluate the directional accuracy of the competing models based on the economic profits from trading implied volatility indices.

Following this naïve trading strategy, we employ a more sophisticated strategy as an additional economic criterion, based on option straddles trading; a straddle is an options strategy in which the investor holds a position in both a call and put option with the same strike price and expiration date. Based on Xekalaki and Degiannakis (2005) and Engle *et al.* (1993) we allow investors to go long (short) in a straddle when the forecasted implied volatility at time $t+s$ is higher (lower) than the actual implied volatility index at the present time *t*. Similar approaches have been employed by Degiannakis and Filis (2017), Andrada-Felix *et al.* (2016), Angelidis and Degiannakis (2008).

The straddle trading is employed given that the straddle holder's rate of return is indifferent to any change in the underlying asset price and is affected only from changes in volatility. Following Engle *et al.* (1993), the next trading day's straddle price on a \$1 share of the underlying stock market index with s days to expiration and \$1 exercise price is:

$$S_{t+1|t} = 4N\left(\frac{\overline{IV}_{t+s|t}}{2}\right) - 2, \tag{10}$$

where $N(.)$ denotes the cumulative normal distribution function and $\overline{IV}_{t+s|t} = \frac{\sum_{i=1}^{s} IV_{t+i|t}}{100\sqrt{252}}$ is the volatility forecast during the life of the option. The daily profit from holding the straddle is $\pi_t = max(e^{y_t} - e^{rf_t}, e^{rf_t} - e^{y_t})$, for y_t denoting the underlying stock market index log-returns and rf_t being the risk-free interest rate.

We assume the existence of thirteen investors who trade their volatility forecasts. Each investor j prices the straddles, $S_{t+1|t}^{(j)}$, every trading day according to one of the thirteen volatility forecasting models[9]. A trade between two investors, j and j^*, is executed at the average of their forecasting prices, yielding to investor j a profit of:

$$\pi_{t+1}^{(j,j^*)} = \begin{cases} \pi_{t+1} - \left(S_{t+1|t}^{(j)} + S_{t+1|t}^{(j^*)}\right) & \text{if } S_{t+1|t}^{(j)} > S_{t+1|t}^{(j^*)} \\ \left(S_{t+1|t}^{(j)} + S_{t+1|t}^{(j^*)}\right) - \pi_{t+1} & \text{if } S_{t+1|t}^{(j)} < S_{t+1|t}^{(j^*)} \end{cases}. \tag{11}$$

As an economic evaluation criterion, we define the cumulative returns computed as $\pi^{(j)} = \frac{1}{2}\sum_{t=1}^{\tilde{T}} \sum_{j^*=1}^{2} \pi_t^{(j,j^*)}$.

6. Empirical findings

6.1. MSE and MAE analysis

We consider the models' forecasting performance at two different horizons, namely 1-day and 10-days ahead. The MSE and MAE loss functions, as well as, the MCS test results are presented in Tables 2 and 3.

[TABLE 2 HERE]

[9] I.e. the HAR, ARFIMA, HW, SSA, SSA-HW, I(1), ARI(1,1), FI(1), ARFI(1,1), ARI-HAR, ARI-ARFIMA, HAR-ARFIMA and ARI-SSA-HW.

[TABLE 3 HERE]

Tables 2 and 3 provide evidence that the forecasts of the SSA-HW model outperform these produced by all naïve, SSA, HW, ARFIMA and HAR models. We observe that this holds true for both time horizons, i.e. 1-day and 10-days ahead, and all indices. The only exception for the 1-day ahead forecasts is the VFTSE, for which the best forecast is achieved by the SSA, according to the MAE. In addition, for the 10-days ahead forecast, the MAE (MSE) suggests that for the VCAC index the best forecast is obtained by the IV-ARFIMA (HW), whereas according to the MSE the best forecasts for the VTFSE and VXD are generated by the HW.

Despite these exceptions, it is clear that the use of the SSA-HW model, as opposed to the naïve, SSA, HW, ARFIMA or HAR models, provides a considerable improvement to the forecasting accuracy for all indices.

Next, we compare the forecasting accuracy of the models using the MCS procedure. The results for the 1-day ahead forecasts (Table 2) suggest that in both the cases of the MAE and the MSE loss functions, the model that belongs to the confident set of the best performing models is only the SSA-HW. The only exception is the forecasts for VFTSE, where in the case of the MAE the best performing model is only the SSA, whereas in the case of MSE it is also the SSA that belongs to the set of the best performing models. For the 10-days ahead forecasts (Table 3), only the SSA-HW is the best one for VXJ and VXN, according to the MSE, whereas for all the other cases, SSA-HW belongs to the set of best models. Based on the MAE, only the SSA-HW is the best model for all the cases except for the VCAC. For the latter, the SSA-HW belongs to the set of the best models.

Overall, evidence suggests that the use of the SSA-HW model offers a substantial improvement to forecasting accuracy, compared to the naïve, SSA, HW, ARFIMA and HAR models.

As a further test for the validity of our findings, we estimate the forecast bias of the SSA-HW relatively to the best performing parametric models (i.e. HAR and ARFIMA). To do so, we employ the Ashley *et al.* (1980) test. We denote as $e_{t+s|t,i} = IV_{t+s|t,i} - IV_{t+s}$ the s-step-ahead forecast error of model *i*, and $\bar{e}_i$ the average of these forecasts. Based on Ashley *et al.* (1980), we are able to estimate the following auxiliary model: $e_{t+s|t,1} - e_{t+s|t,2} = a + b\left(\left(e_{t+s|t,1} + e_{t+s|t,2}\right) - \left(\bar{e}_1 + \bar{e}_2\right)\right) + z_{t+s}$, for

$z_t \sim N\left(0, \sigma_z^2\right)$. A statistically significant intercept provides evidence that there is significant difference in the forecast errors. Moreover, a statistically significant slope shows a difference in the forecast error variances. Overall, we may investigate the null hypothesis that the difference between the two forecasting models is statistically negligible. As Ashley *et al.* (1980) noted, in the case that either of the two least squares estimates is significantly negative, the model (1) (i.e. SSA-HW in our case) provides superior forecasts[10]. The results are reported in Table 4.

[TABLE 4 HERE]

From Table 4 we find evidence that the improvement in the forecasts of the implied volatilities using the SSA-HW model primarily stems from the reduction in the variance of the forecast errors, given that the *b* coefficient is negative and significant, relatively to the HAR and ARFIMA models.

6.2. SSA-HW performance over time

The aforementioned results provide a convincing picture that the SSA-HW is the best performing forecasting model for both the 1-day and 10-days ahead horizons. Next we evaluate whether its predictive ability holds during different market conditions, namely, during periods characterized by high or low volatility. To do so, we calculate the incremental predictive ability of the SSA-HW model relatively to the best performing parametric models, i.e. HAR and ARFIMA. Motivated by Degiannakis and Filis (2017), the incremental value of the SSA-HW is captured by the cumulative difference between its MAE relatively to the MAE of the HAR and ARFIMA models, separately. Figures 2 and 3 depict these cumulative differences for the 1-day and 10-days ahead horizons, respectively.

[FIGURE 2 HERE]

[FIGURE 3 HERE]

We should note that when the cumulative difference increases then the SSA-HW exhibits incremental predictive gains, whereas the reverse holds true with the cumulative difference decreases. Figures 2 and 3 reveal that in almost all cases the SSA-HW does provide incremental predictive gains compared to the two best

[10] If one estimate is negative and statistically insignificant, then a one-tailed t-test on the other coefficient can be used. If both estimates are positive, an F test for the null hypothesis that both coefficients are statistically zero can be applied (half of the significance level reported from the tables must be reported).

performing parametric models, i.e. the HAR and ARFIMA (although this does not apply to the post-global financial crisis for the 1-day ahead horizon of VFTSE and the 10-days ahead horizon of VCAC). It is also important to highlight that almost all figures exhibit a steeper increase during the 2008-09 period, i.e. the global financial crisis. This is suggestive of the fact that during turbulent times the SSA-HW provides even higher incremental predictive gains.

The last observation even holds for the case of the 10-days ahead forecast of the VCAC, for which we documented that the SSA-HW does not provide the most accurate forecasts. More specifically, a steep upward movement in the VCAC figure is observed during the global financial crisis, suggesting that for this period the SSA-HW does provide very high incremental predictive gains relatively to the HAR and ARFIMA models.

This is further evidence that SSA-HW not only exhibits a high forecasting ability, but also its ability is stronger during turbulent times, when accurate forecasts are even more necessary.

6.3. Model-averaged forecasts

Next, we proceed with model-averaged forecasts in order to assess whether the inclusion of a naïve model could improve the performance of the competing models. According to Tables 2 and 3 the best naïve model is the ARI(1,1) model. Thus, we consider the following model-averaged forecasts, ARI-IV-ARFIMA, ARI-IV-HAR and ARI-SSA-HW. In addition, we also use the model-averaged forecast of the ARFIMA-HAR models. Table 5 summarizes the results for the 1-day and 10-days ahead forecasts for both the MSE and the MAE.

[TABLE 5 HERE]

For the 1-day ahead forecasts, we observe that apart from the VCAC, VDAX and VSTOXX, in all other cases the model-averaged forecasts based on the ARI-SSA-HW can outperform the SSA-HW. Even more, for the 10-days ahead forecasts, we notice that the inclusion of the ARI(1,1) model in the SSA-HW is able to produce superior predictions for all implied volatility indices.

To assess further the superior predictive ability of the ARI-SSA-HW, we perform the MCS test including all competing models, i.e. the original nine models, as well as, the model-averaged forecasts. For brevity, Table 6 presents the MCS *p*-values of the best performing models only, for the 1-day ahead and 10-days ahead horizons.

[TABLE 6 HERE]

Table 6 suggests that for the 1-day ahead forecasts, in almost all cases the SSA-HW model belongs to the set of the best performing models along with the ARI-SSA-HW. The only exception is the VXJ, where only the ARI-SSA-HW is included in the set of the best performing models. Thus, even though the model-averaged forecasts improve the forecasting accuracy of the SSA-HW model, this improvement is not significantly higher for all implied volatility indices.

The MCS results for the 10-days ahead forecasts (see Table 6) reveal that the ARI-SSA-HW model is always among the best performing models; yet, the SSA-HW also belongs to the set of the best models in three cases (VDAX, VFTSE and VIX). HW is also among the best models for the case of VFTSE. Thus, our study presents empirical evidence that in the case of multi-days-ahead volatility forecasts the predictive accuracy of the model-averaged method is statistically significantly improved.

Scatter plots in Figure 4 provide a visual representation of the relationship between actual and predicted implied volatility indices for the VIX index, indicatively. Panel A corresponds to the 1-day ahead forecasts, whereas Panel B exhibits the 10-days ahead forecasts. These scatter plots rendered it clear that the SSA-HW produces the slimmest plots (middle column) for the 1-day ahead forecast, whereas for the 10-days ahead forecast it is the ARI-SSA-HW (right column). The worse forecasts are produced by the FI(1,1) for both forecasting horizons. In addition, the SSA-HW for the 1-day ahead and the ARI-SSA-HW model for the 10-days ahead forecasts are observed to have fewer outliers. In addition, it is worth noting that at the higher levels of volatility, the SSA-HW (for the 1-day ahead) and the ARI-SSA-HW (for the 10-days ahead) models appear to produce less scattered points.

[FIGURE 4 HERE]

Overall, the SSA-HW model, along with the ARI-SSA-HW, are superior to their competitors, for the 1-day ahead forecast, whereas the combination of SSA-HW with the ARI(1,1) is the best model for the 10-days ahead. We also assess the forecasting performance of our models in three sub-periods (pre-crisis period: January 2005 – November 2007, crisis period: December 2007 – June 2009, post-crisis period:

July 2009 – July 2013) and the results are qualitatively similar. For brevity, these results are available upon request.

The ability of the SSA-HW to generate superior forecasts stems from the fact that it utilises the advantages of each of the model's components. The SSA has the ability to decompose volatility indices into interpretable components. By decomposing the series using SSA, the interpretable components capture the dynamics of volatility indices, which can then be forecasted individually using HW. In turn, HW can provide accurate forecasts of trend and signal via exponentially weighted moving averages (Holt, 2004). Thus, HW's modelling capability is enhanced by the SSA filtering, which reduces the noise of the series. Therefore, instead of forecasting the index itself, we forecast each decomposed series prior to combining these forecasts.

In more simple terms, the superior performance reported by SSA-HW can be attributed to the fact that in the absence of filtering with SSA, the trend and other signals within the index would be distorted by the noise. When we decompose the series, we are able to separate all such components into individual time series where each series will have its own and varying structure, earlier hidden underneath the overall series. Thereby, forecasting these individual series (extracted from SSA) with HW enables us to capture the underlying fluctuations, which would have been more difficult to reveal without SSA filtering. This is further evidenced by the fact that neither SSA nor HW is able to outperform the forecasts of SSA-HW at both horizons, apart from few exceptions.

Furthermore, SSA is more popular as a filtering technique as opposed to a forecasting technique. This might explain its poor forecasting performance, as the SSA forecasting algorithm appears to encounter problems with modelling implied volatility even after filtering for noise. Note that when SSA filters for noise, it forecasts the signal alone and, contrary to the SSA-HW approach, this is not decomposed further. Similarly, HW's poor predictive performance is attributable to the fact that there is no filtering involved and as a result, it encounters problems in identifying the true signal, which is distorted by the noise component of the implied volatility indices.

6.4. Direction of change

The DoC results are shown in Tables 7 and 8 for the 1-day and 10-days ahead, respectively. Table 7 shows that all forecasting models exhibit a good prediction of the DoC, since all scores are above the 50% level (with the only exception being the I(1) model), nevertheless the forecasting model with the highest prediction ability is the SSA-HW, followed by the ARI-SSA-HW and the SSA. More specifically, the SSA-HW and ARI-SSA-HW are capable of predicting the DoC accurately in 65-80% of the cases, depending on the volatility index. Similar findings are reported for the 10-days ahead forecasts (as shown in Table 8), where the SSA-HW and ARI-SSA-HW exhibit a very high predictive ability of the DoC, although the highest precision is attributed to the SSA-HW. In particular, the models are able to predict 65-88% of the directional changes of the implied volatilities. These results corroborate the findings of the MCS, which provided evidence that the best model is the SSA-HW, followed by the ARI-SSA-HW.

[TABLES 7 and 8 HERE]

6.5. Forecasting performance based on the trading strategies

The results of the trading strategy are reported in Tables 9 and 10 for the 1-day and 10-days ahead, respectively.

[TABLES 9 and 10 HERE]

For the 1-day ahead (see Table 9), it is evident that the SSA, SSA-HW and the ARI-SSA-HW provide positive net returns, which are significantly higher than zero. The largest figures are observed for the SSA-HW, followed by the ARI-SSA-HW and the SSA. Turning our attention to the 10-days ahead (see Table 10), we can make a similar inference, as the only forecasting models that yield positive net returns are those of the HW, SSA-HW and ARI-SSA-HW. Nevertheless, we observe that statistically significant net returns are only feasible for the VIX and VSTOXX indices. Hence, these findings confirm the superior predictive ability of the SSA-HW.

Finally, Tables 11 and 12 present the cumulative returns of investors who are pricing their straddles according to the implied volatility forecasts from the thirteen competing models. The results show that the SSA-HW and the ARI-SSA-HW models are able to generate superior positive profits against the other competing models, although this does not apply to all implied volatility indices. We should highlight here that even when investors, who use the aforementioned models, do not obtain the highest positive profits, their trading strategies are in almost all cases among the most profitable. In any case, the option straddles trading strategy provides some additional evidence that the SSA-HW and the ARI-SSA-HW are capable of producing forecasts that are economically important.

[TABLES 11 and 12 HERE]

7. Conclusion

The aim of this paper is to compare parametric and non-parametric techniques in terms of their forecasting power for implied volatility indices. We extend our comparisons using combined and model-averaging models. More specifically, we generate 1-day and 10-days ahead forecasts based on the SSA, HW, ARFIMA and HAR models, as well as, combined models and model-averaged frameworks. In addition, we use four naïve models. We compare their forecasting accuracy using the MSE and MAE evaluation criteria, the MCS procedure and the Direction-of-Change. Furthermore, we assess the forecasting ability of the models using two trading strategies.

The results show that the SSA-HW is a powerful tool for predicting implied volatility indices as it is able to exploit the advantages of two non-parametric methods. The forecasting accuracy tests reveal that the forecasts generated by the SSA-HW model outperform these by the naïve, ARFIMA and HAR models for the 1-day ahead. On the other hand, the model-averaged forecasts reveal that the ARI-SSA-HW improves the SSA-HW forecasts, particularly for the 10-days ahead forecasts.

The results of the trading strategies confirm these findings, revealing that the SSA-HW and the ARI-SSA-HW could provide significantly positive net returns over the out-of-sample period, although this primarily holds for the 1-day ahead. Overall, we maintain that this superior forecasting ability of the non-parametric techniques, as well as, the model-averaging between parametric and non-parametric model is important to investors (e.g. for portfolio allocation decisions), portfolio managers (e.g. for Global Tactical Asset Allocation strategies), derivatives pricing, risk management purposes, as well as, policy makers (e.g. monetary policy decisions).

The use of SSA-HW enables us to overcome the parametric assumptions, which restrict the applicability of many parametric models to real world scenarios. As such we believe this proposed forecasting framework, which combines a renowned forecasting technique (HW) with an equally renowned filtering technique (SSA), will enable users to achieve better outcomes when applied to other real world forecasting problems, which go beyond implied volatility forecasts. In a world where the emergence of Big Data and the related noise continue to distort the signal in time series, the proposed SSA-HW approach can be a useful tool for attaining reliable and accurate forecasts in the future. An interesting avenue for further study is to assess SSA forecasting ability using intra-day data.

Acknowledgements

The authors would like thank the editor (Prof. Rossen I. Valkanov), the associate editor and the anonymous referee for their invaluable comments and suggestions on a previous version of this paper, which helped us to improve significantly the quality of the paper. The usual disclaimer applies for any remaining errors and omissions.

References

Alexandrov, T. (2009). A method of trend extraction using singular spectrum analysis. *REVSTAT,* 7(1), 1-22.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), 579-625.

Andersen, T. G., Bollerslev, T., and Meddahi, N. (2005). Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities. *Econometrica*, 73(1), 279-296.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Review of Economics and Statistics*, 89, 701–720.

Angelidis, T., Benos, A. and Degiannakis, S. (2004). The Use of GARCH Models in VaR Estimation, *Statistical Methodology*, 1(2), 105-128.

Angelidis, T., and Degiannakis, S. (2008). Volatility forecasting: Intra-day versus inter-day models. *Journal of International Financial Markets, Institutions and Money*, 18(5), 449-465.

Barunik, J., Krehlik, T., and Vacha, L. (2016). Modeling and forecasting exchange rate volatility in time-frequency domain. *European Journal of Operational Research*, 251(1), 329-340.

Bårdsen, G. and Lütkepohl, H. (2011). Forecasting levels of log variables in vector autoregressions. *International Journal of Forecasting*, 27(4), 1108-1115.

Beckers, S. (1981). Standard deviations implied in options prices as predictors of future stock price variability. *Journal of Banking and Finance,* 5, 363– 381.

Beneki, C., Eeckels, B., and Leon, C. (2012). Signal Extraction and Forecasting of the UK Tourism Income Time Series: A Singular Spectrum Analysis Approach. *Journal of Forecasting*, 31(5), 391-400.

Bildirici, M., and Ersin, Ö. Ö. (2009). Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange. *Expert Systems with Applications*, 36(4), 7355-7362.

Blair, B.J., Poon, S-H and Taylor S.J. (2001). Forecasting S&P100 Volatility: The Incremental Information Content of Implied Volatilities and High-Frequency Index Returns. *Journal of Econometrics*, 105, 5-26.

Bollerslev, T., Engle, R. F. and Nelson, D. (1994). ARCH models, in *Handbook of Econometrics*, Vol. 4 (Eds) R. F. Engle and D. McFadden, Elsevier Science, Amsterdam, 2959–3038.

Booth, H., and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science,* 3(1-2), 3-43.

Busch, T., Christensen, B. J., and Nielsen, M. Ø. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics*, 160, 48–57.

Busemeyer, J.R., Wang, Z., Townsend, J.T., and Eidels, A. (2015). *The Oxford Handbook of Computational and Mathematical Psychology.* Oxford University Press.

Charles, A. (2010). The day-of-the-week effects on the volatility: The role of the asymmetry. *European Journal of Operational Research*, 202(1), 143-152.

Chatfield, C. (1996). Model uncertainty and forecast accuracy. *Journal of Forecasting,* 15(7), 495-508.

Chiras, D.P., and Manaster, S. (1978). The information content of options prices and a test of market efficiency. *Journal of Financial Economics*, 6, 213–234.

Christensen, B.J., and Prabhala, N.R. (1998). The relation between implied and realised volatility. *Journal of Financial Economics*, 50, 125– 150.

Christodoulakis, G. A. (2007). Common volatility and correlation clustering in asset returns. *European Journal of Operational Research*, 182(3), 1263-1284.

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

Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), 174-196.

Corsi, F., Kretschmer, U., Mittnik, S. and Pigorsch, C. (2005). The Volatility of Realised Volatility. Center for Financial Studies, Working Paper, 33

Degiannakis, S. (2004). Volatility forecasting: evidence from a fractional integrated asymmetric power ARCH skewed-t model. *Applied Financial Economics*, 14, 1333–1342.

Degiannakis, S. (2008a). Forecasting VIX. *Journal of Money, Investment and Banking*, 4, 5-19.

Degiannakis, S. (2008b). ARFIMAX and ARFIMAX- TARCH Realized Volatility Modeling. *Journal of Applied Statistics*, 35(10), 1169-1180.

Degiannakis, S., and Filis, G. (2017). Forecasting oil price realized volatility using information channels from other asset classes. *Journal of International Money and Finance*, 76, 28-49.

Degiannakis, S., Livada, A. and Panas, E. (2008). Rolling-sampled parameters of ARCH and Levy-stable models. *Journal of Applied Economics*, 40(23), 3051-3067.

Deo, R., Hurvich, C., and Lu, Y. (2006). Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment. *Journal of Econometrics*, 131(1), 29-58.

Diebold, F. X., and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13, 253-263.

Donaldson, R. G., and Kamstra, M. (1997). An artificial neural network-GARCH model for international stock return volatility. *Journal of Empirical Finance*, 4(1), 17-46.

Doornik, J.A. and Ooms, M. (2006). A Package for Estimating, Forecasting and Simulating Arfima Models: Arfima Package 1.04 for Ox. Working Paper, Nuffield College, Oxford.

Engle, R.F., Hong, C.H., Kane, A. and Noh, J. (1993). Arbitrage Valuation of Variance Forecasts with Simulated Options, *Advances in Futures and Options Research,* 6, 393-415.

Favero, C. and Aiolfi, M. (2005). Model uncertainty, thick modelling and the predictability of stock returns, *Journal of Forecasting*, 24, 233-254.

Fernandes, M., Medeiros, M. C., and Scharth, M. (2014). Modeling and predicting the CBOE market volatility index. *Journal of Banking & Finance*, 40, 1-10.

Figlewski, S. (1997). Forecasting volatility. *Financial markets, Institutions & Instruments*, 6(1), 1-88.

Fleming, J. (1998). The quality of market volatility forecast implied by S&P 100 index option prices. *Journal of Empirical Finance*, 5, 317– 345.

Fleming, J., Ostdiek, B. and Whaley, R.E. (1995). Predicting Stock Market Volatility: A New Measure. *Journal of Futures Markets*, 15, 265-302.

Frijns, B., Tallau, C., and Tourani-Rad, A. (2010). The information content of implied volatility: evidence from Australia. *Journal of Futures Markets*, 30(2), 134-155.

Fuertes, A. M., Izzeldin, M., & Kalotychou, E. (2009). On forecasting daily stock volatility: The role of intraday information and market conditions. *International Journal of Forecasting*, 25(2), 259-281.

Giot, P. (2003). The information content of implied volatility in agricultural commodity markets. *Journal of Futures Markets*, 23, 441–454.

Golyandina, N., Korobeynikov, A., Shlemov, A., and Usevich, K. (2015). Multivariate and 2D extensions of singular spectrum analysis with the Rssa package. *Journal of Statistical Software*, 67(2), DOI: 10.18637/jss.v067.i02.

Granger, C.W.J. and Newbold, P. (1976). Forecasting transformed series. *Journal of the Royal Statistical Society, Series B*, 38, 189–203.

Granger, C.W.J. and Joyeux, R. (1980). An Introduction to Long Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, 1, 15-39.

Granger, C. W., and Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3(2), 197-204.

Hajizadeh, E., Seifi, A., Zarandi, M. F., and Turksen, I. B. (2012). A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, 39(1), 431-436.

Hansen, P.R., 2005. A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23, 365–380.

Hansen, P.R., and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1,1)? *Journal of Applied Econometrics*, 20(7), 873-889.

Hansen, P.R., Lunde, A. and Nason, J.M. (2011). The model confidence set, *Econometrica*, 79, 456–497.

Harvey, A. (1990). *The econometric analysis of time series.* Philip Allan, London.

Hassani, H., Heravi, S., and Zhigljavsky, A. (2009). Forecasting European Industrial Production with Singular Spectrum Analysis. *International Journal of Forecasting,* 25(1), 103-118.

Hassani, H., and Zhigljavsky, A. (2009). Singular spectrum analysis: methodology and application to economics data. *Journal of System Science and Complexity*, 223, 372–394.

Hassani, H., and Thomakos, D. (2010). A review on Singular Spectrum Analysis for economic and financial time series. *Statistics and Its Interface*, 3, 377-397.

Hassani, H., **Heravi, S., Brown, G., and Ayoubkhani, D. (2013).** Forecasting before, during, and after recession with Singular Spectrum Analysis. *Journal of Applied Statistics,* 40(10), 2290-2302.

Hassani, H., Silva, E.S., Antonakakis, N., Filis, G., and Gupta, R. (2017). Forecasting accuracy evaluation of tourist arrivals. *Annals of Tourism Research,* 63, 112-127.

Holt, C. C. (2004). Forecasting trends and seasonals by exponentially weighted moving averages. *International Journal of Forecasting,* 20(1), 5-10.

Hung, J. C. (2011a). Applying a combined fuzzy systems and GARCH model to adaptively forecast stock market volatility. *Applied Soft Computing*, 11(5), 3938-3945.

Hung, J. C. (2011b). Adaptive Fuzzy-GARCH model applied to forecasting the volatility of stock markets using particle swarm optimization. *Information Sciences*, 181(20), 4673-4683.

Hyndman, R. J., and Athanasopoulos, G. (2014). Forecasting: principles and practice. OTexts**.**

Hyndman, R. J., Athanasopoulos, G., Razbash, S., Schmidt, D., Zhou, Z., Khan, Y., and Bergmeir, C. (2013). Package forecast: Forecasting functions for time series and linear models. Available via: http://cran.r-project.org/web/packages/forecast/forecast.pdf

Jung, Y. C. (2016). A portfolio insurance strategy for volatility index (VIX) futures. *The Quarterly Review of Economics and Finance*, 60, 189-200.

Koopman, S.J., Jungbacker, B., and Hol, E. (2005). Forecasting daily variability of the S&P100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance*, 12(3), 445–475.

Kristjanpoller, W., Fadic, A., and Minutolo, M. C. (2014). Volatility forecast using hybrid neural network models. *Expert Systems with Applications*, 41(5), 2437-2442.

McLeod, A.I. (1993). Parsimony, Model adequacy and Periodic Correlation in Time Series Forecasting. *International Statistical Review,* 61(3), 387-393.

Mincer, J. and Zarnowitz, V. (1969). The Evaluation of Economic Forecasts. In (ed.) Mincer. J., *Economic Forecasts and Expectations*, National Bureau of Economic Research, , New York, 3-46.

Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V. and VonWeizsäcker, J.E. (1997). Volatilities of Different Time Resolutions – Analyzing the Dynamics of Market Components. *Journal of Empirical Finance*, 4, 213-239.

Ortu, F., Tamoni, A., and Tebaldi, C. (2013). Long-run risk and the persistence of consumption shocks. The Review of Financial Studies, 26(11), 2876-2915.

Samuels, J.D., and Sekkel, R.M. (2013). Forecasting with Many Models: Model Confidence Sets and Forecast Combination. Working Paper, Bank of Canada.

Silva, E.S., and Hassani, H. (2015). On the use of Singular Spectrum Analysis for forecasting U.S. trade before, during and after the 2008 recession. *International Economics,* 141, 34-49.

Soofi, A., and Cao, L.Y. (2002). *Nonlinear forecasting of noisy financial data: Techniques of Nonlinear Dynamics.* Kluwer Academic Publishers, Boston.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6, 461-464.

Sevi, B. (2014). Forecasting the Volatility of Crude Oil Futures using Intraday Data. *European Journal of Operational Research*, 235, 643-659.

Simon, D. P. (2003). The Nasdaq volatility index during and after the bubble. *The Journal of Derivatives*, 11, 9–24.

Thomakos, D. D., Wang, T., and Wille, L. T. (2002). Modeling daily realized futures volatility with singular spectrum analysis. *Physica A: Statistical Mechanics and its Applications*, 312(3), 505-519.

Timmermann, A. (2006). Forecast combinations. *Handbook of Economic Forecasting*, 1, 135-196.

Vautard, R., Yiou, P., and Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenomena*, 58(1), 95-126.

Wei, L. Y. (2013). A GA-weighted ANFIS model based on multiple stock market volatility causality for TAIEX forecasting. *Applied Soft Computing*, 13(2), 911-920.

White, H. (2000). A Reality Check for Data Snooping. *Econometrica*, 68, 1097–1126.

Xekalaki, E. and Degiannakis, S. (2010). *ARCH models for financial applications*. Wiley and Sons, New York.

Yu, J. (2012). A semiparametric stochastic volatility model. *Journal of Econometrics*, 167(2), 473-482.

Figures

Figure 1: Implied Volatility Indices. The sample period runs from January, 2001 to July, 2013.

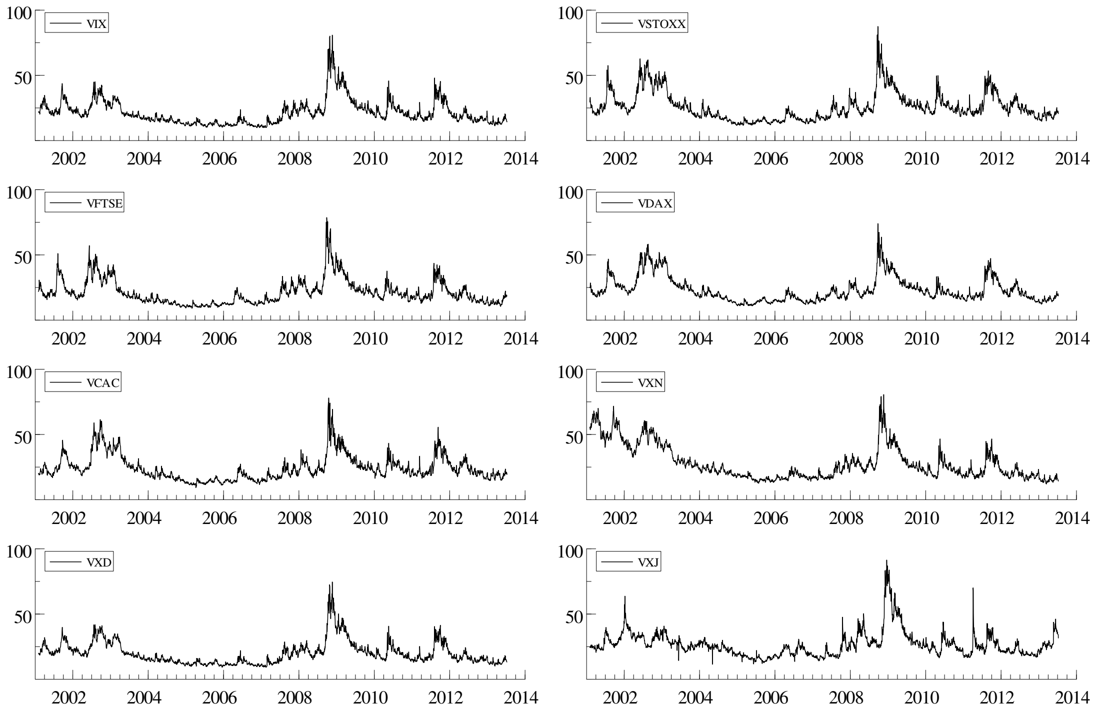
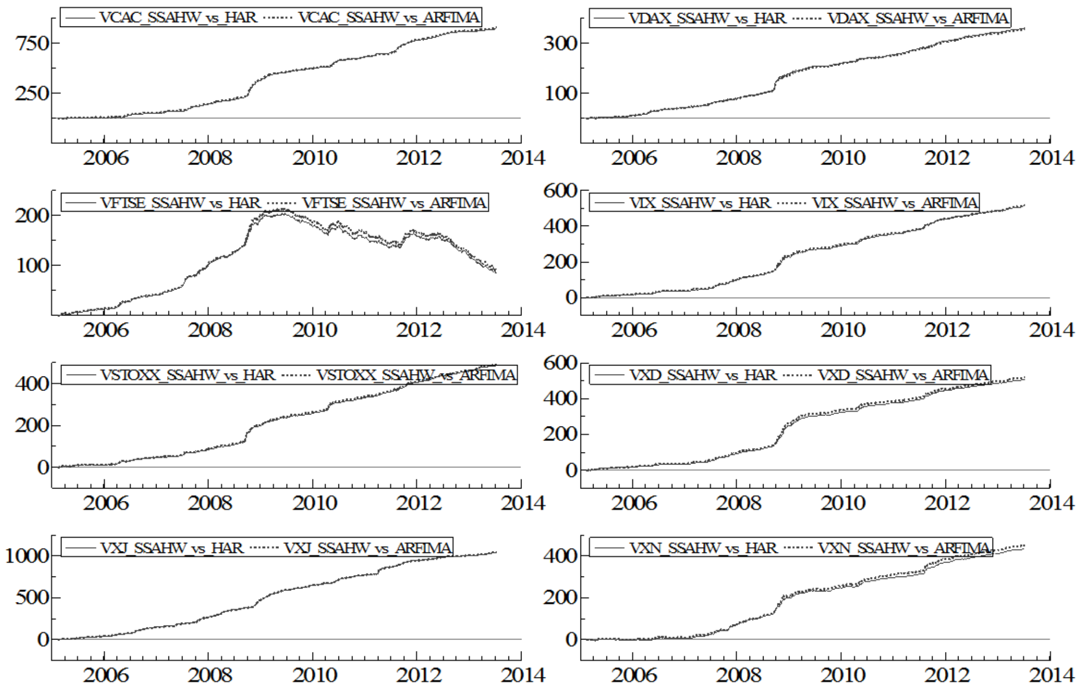
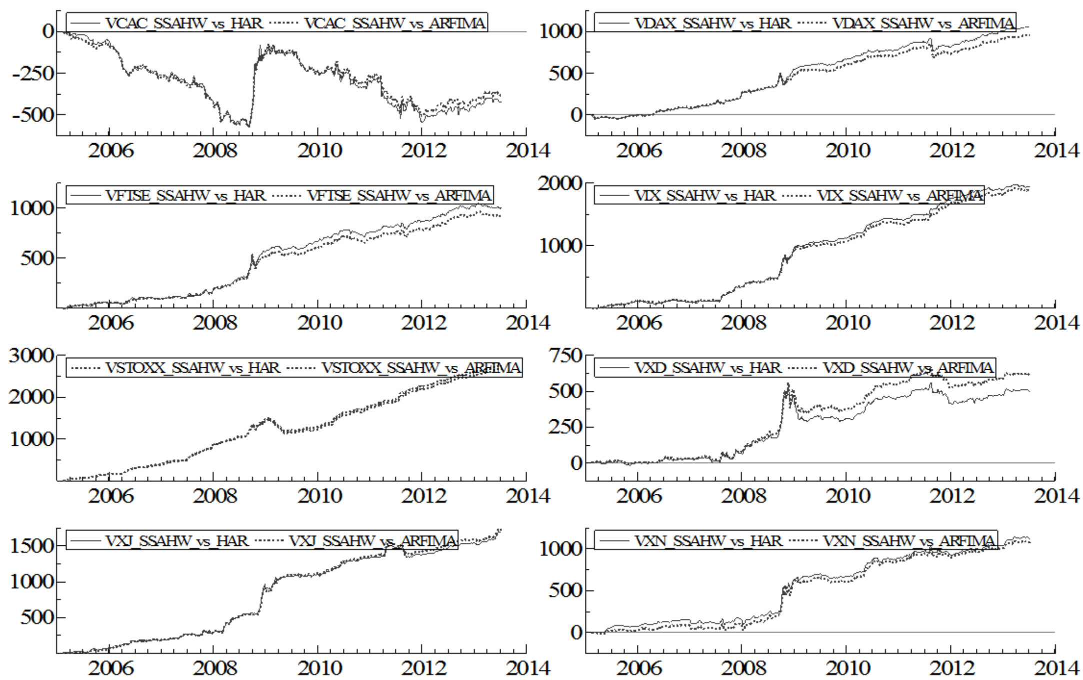


Figure 2: Cumulative incremental predictive gains of the IV-SSA-HW model vs. the IV-HAR and IV-ARFIMA for the 1-day ahead, based on the MAE.



Note: Upward (downward) movements suggest that the IV-SSA-HW (IV-HAR or IV-ARFIMA) provides the best predictive gains.

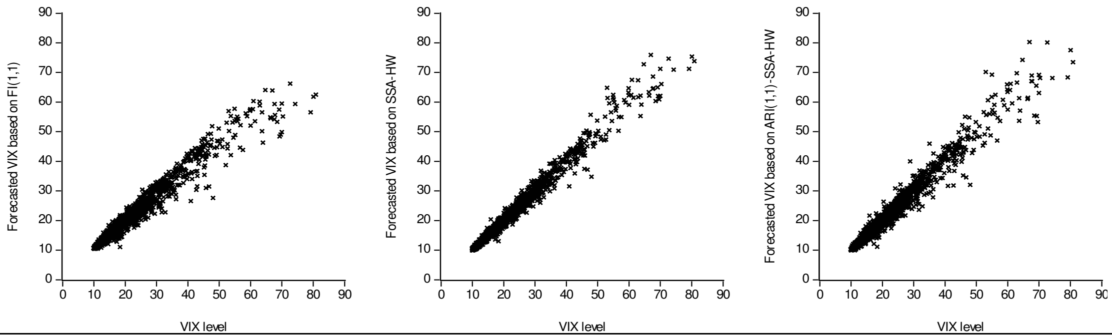
Figure 3: Cumulative incremental predictive gains of the IV-SSA-HW model vs. the IV-HAR and IV-ARFIMA for the 10-days ahead, based on the MAE.



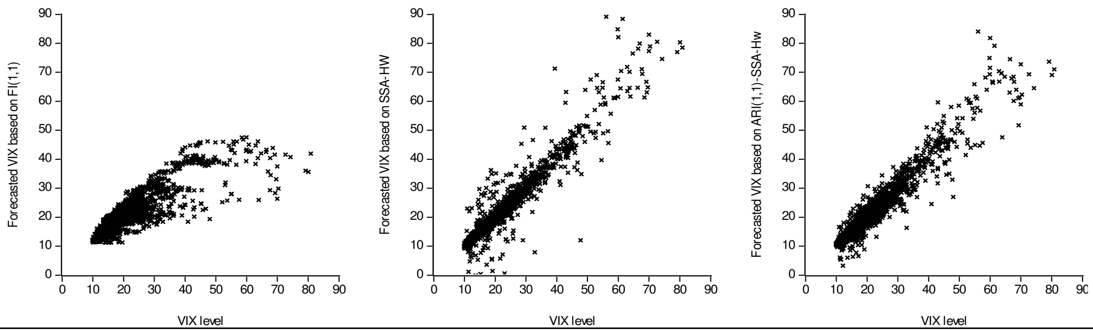
Note: Upward (downward) movements suggest that the IV-SSA-HW (IV-HAR or IV-ARFIMA) provides the best predictive gains.

Figure 4: One-day and 10-days ahead forecasts scatter plots of the models for the VIX index. The sample period runs from January, 2005 to July, 2013.

1-day ahead forecasts



10-days ahead forecasts



Note: Columns from left to right present the scatter plots for FI(1,1), SSA-HW and ARI(1,1)-SSA-HW, respectively. The y-axes (x-axes) show the actual (predicted) values.

Tables

Table 1: Descriptive Statistics of Implied Volatility Indices (January, 2001 to July, 2013).

	Mean	Min	Max	Std.Dev	Jarque-Bera	ADF-statistic	ARCH LM Test
VIX	21.52	9.89	80.86	9.48	6174.43 ***	-3.23 **	5288.04 ***
VSTOXX	25.99	11.60	87.51	10.78	1655.11 ***	-3.63 ***	5759.33 ***
VFTSE	21.19	9.10	78.69	9.45	3829.52 ***	-3.89 ***	5535.42 ***
VDAX	23.32	10.98	74.00	9.54	1578.59 ***	-3.16 **	8317.23 ***
VCAC	24.31	9.24	78.05	9.76	2250.23 ***	-3.69 ***	4588.81 ***
VXN	27.92	12.03	80.64	13.01	929.13 ***	-2.98 **	12370.04 ***
VXD	19.98	9.28	74.60	8.80	5205.14 ***	-3.17 **	6263.71 ***
VXJ	26.66	11.53	91.45	9.70	12706.03 ***	-4.10 ***	5620.22 ***

***,**,* indicate significance at 1%, 5% and 10% level, respectively.

Table 2: Forecast accuracy tests: One-day ahead forecasts (January, 2005 to July, 2013).

		Implied Volatility Indices							
Model	Loss Function	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
IV-HAR	MSE	4.18	2.21	2.92	3.81	3.76	2.91	4.67	3.12
	MAE	1.21	0.90	1.06	1.15	1.17	1.03	1.24	1.10
IV-ARFIMA	MSE	4.20	2.19	2.90	3.84	3.77	2.96	4.67	3.18
	MAE	1.22	0.90	1.06	1.16	1.17	1.04	1.25	1.10
HW	MSE	4.65	2.76	3.54	4.42	4.90	3.36	5.46	4.18
	MAE	1.37	1.11	1.28	1.34	1.49	1.19	1.45	1.44
SSA	MSE	2.55	1.67	2.39*	2.92	2.87	2.09	2.71	2.41
	MAE	0.99	0.81	**0.98***	1.04	1.05	0.91	0.97	0.99
SSA-HW	MSE	**1.46***	**1.29***	**2.28***	**2.18***	**2.20***	**1.49***	**1.46***	**1.86***
	MAE	**0.79***	**0.73***	1.02	**0.91***	**0.94***	**0.79***	**0.75***	**0.89***
I(1)	MSE	4.28	2.21	2.94	3.96	3.81	3.00	4.64	3.16
	MAE	1.22	0.90	1.06	1.16	1.18	1.04	1.24	1.10
ARI(1,1)	MSE	4.26	2.22	2.93	3.86	3.81	2.94	4.70	3.15
	MAE	1.22	0.90	1.06	1.16	1.18	1.03	1.25	1.10
FI(1)	MSE	6.11	3.98	5.23	6.07	6.29	4.75	8.22	5.20
	MAE	1.45	1.17	1.32	1.39	1.45	1.26	1.54	1.35
ARFI(1,1)	MSE	4.37	2.33	3.10	4.28	3.96	3.27	5.14	3.42
	MAE	1.24	0.92	1.07	1.19	1.18	1.06	1.30	1.13

Bold face fonts present the models with the lowest values of MAE and MSE. * denotes that the model is included in the set of the best performing models, according to the MCS test.

Table 3: Forecast accuracy tests: Ten-days ahead forecasts (January, 2005 to July, 2013).

Model	Loss Function	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
		Implied Volatility Indices							
IV-HAR	MSE	21.22*	13.86*	19.85	18.94	22.17	15.60*	29.57	18.88
	MAE	2.92*	2.39	2.77	2.72	2.96	2.50	3.20	2.74
IV-ARFIMA	MSE	21.27*	13.47*	19.41	19.32	21.89	15.56*	29.18	19.61
	MAE	**2.90***	2.34	2.73	2.69	2.93	2.44	3.19	2.76
HW	MSE	**17.77***	13.36*	**14.04***	13.98*	19.03*	**13.51***	21.90	18.66
	MAE	2.91*	2.27	2.38	2.38	2.73	2.49	2.74	3.04
SSA	MSE	45.80	19.78	33.12	26.24	36.10	24.52	54.05	34.66
	MAE	4.26	2.72	3.46	3.20	3.58	3.22	4.32	3.69
SSA-HW	MSE	20.41*	**12.12***	14.99*	**13.13***	**15.49***	14.40*	**19.00***	**12.70***
	MAE	3.10*	**1.89***	**2.29***	**1.79***	**1.66***	**2.21***	**2.39***	**2.22***
I(1)	MSE	22.22	13.77*	20.15	18.56	22.56	14.93*	30.19	18.37
	MAE	3.05	2.42	2.83	2.74	3.08	2.50	3.26	2.77
ARI(1,1)	MSE	21.98	13.75*	20.11	18.35	22.49	14.81*	30.11	18.29
	MAE	3.03	2.42	2.83	2.74	3.08	2.50	3.25	2.77
FI(1)	MSE	28.12	21.69	27.82	31.20	32.24	25.22	42.89	27.89
	MAE	3.21	2.82	3.10	3.23	3.38	2.93	3.78	3.22
ARFI(1,1)	MSE	26.55	19.69	25.65	29.43	29.84	23.72	41.37	26.03
	MAE	3.11	2.67	2.97	3.13	3.25	2.84	3.69	3.09

Bold face fonts present the models with the lowest values of MAE and MSE. * denotes that the model is included in the set of the best performing models, according to the MCS test.

Table 4: Bias and forecast variance reduction for the 1-day and 10-days ahead horizons (January, 2005 to July, 2013).

		Implied Volatility Indices							
Forecast horizon	Coeff.	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
		SSA-HW vs HAR							
1-day ahead	a	-0.003	-0.001	-0.000	-0.013	0.006	-0.021	0.008	-0.044
	b	-0.286***	-0.144***	-0.068***	-0.146***	-0.142***	-0.178***	-0.307***	-0.137***
10-days ahead	a	0.191	0.121	0.049	0.111	0.192	0.178	0.202	0.209
	b	-0.014	-0044***	-0.085***	-0.174***	-0.214***	-0.026*	-0.156***	-0.166***
		SSA-HW vs ARFIMA							
1-day ahead	a	0.010	0.011	0.018	0.011	0.025	0.010	0.016	0.020
	b	-0.289***	-0.143***	-0.066***	-0.148***	-0.143***	-0.183***	-0.307***	-0.142***
10-days ahead	a	0.117	0.226	0.196	0.265	0.155	-0.049	0.115	0.121
	b	-0.014	-0.034***	-0.078***	-0.188***	-0.205***	-0.028**	-0.161***	-0.151***

Note: *, **, *** denote significance at the 10%, 5% and 1% level, respectively. The a and b coefficients are estimated based on the Ashley *et al.* (1980) auxiliary regression model. If the a coefficient is negative and significant then it denotes a forecast bias reduction of the SSA-HW relatively to the HAR and ARFIMA models. If the b coefficient is negative and significant this denotes a reduction in the forecast error variance from the SSA-HW relatively to the other two models.

Table 5: Forecast accuracy tests: Model-averaged forecasts based on the Granger and Ramanathan (1985) approach (January, 2005 to July, 2013).

		Implied Volatility Indices							
Model	Loss Function	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
		One-day ahead							
ARI-IV-HAR	MSE	4.11	2.39	3.06	4.06	4.10	3.07	4.82	3.28
	MAE	1.19	0.92	1.08	1.17	1.20	1.05	1.26	1.11
ARI-IV-ARFIMA	MSE	4.15	2.40	3.08	4.15	4.06	3.11	4.85	3.31
	MAE	1.20	0.92	1.08	1.17	1.20	1.05	1.27	1.11
HAR-ARFIMA	MSE	4.48	2.37	3.09	4.07	4.02	3.08	4.84	3.26
	MAE	1.24	0.92	1.09	1.18	1.20	1.06	1.28	1.12
ARI-SSA-HW	MSE	1.51	1.32	**2.20**	**2.06**	2.24	**1.39**	**1.20**	**1.83**
	MAE	0.79	0.73	**0.94**	**0.89**	0.95	**0.78**	**0.71**	**0.88**
		Ten-days ahead							
ARI-IV-HAR	MSE	21.01	13.02	18.84	17.41	21.44	14.48	27.42	16.56
	MAE	2.92	2.36	2.72	2.65	2.94	2.47	3.25	2.64
ARI-IV-ARFIMA	MSE	21.18	13.33	19.67	17.07	21.55	14.02	27.80	16.61
	MAE	2.95	1.67	2.81	2.63	2.96	2.43	3.25	2.63
HAR-ARFIMA	MSE	22.24	14.13	20.07	20.19	22.88	17.36	29.38	19.46
	MAE	3.02	2.45	2.87	2.82	3.09	2.62	3.43	2.84
ARI-SSA-HW	MSE	**13.64**	**9.68**	**13.79**	**7.83**	**7.98**	**10.77**	**15.52**	**10.07**
	MAE	**2.45**	1.92	**2.27**	**1.72**	**1.55**	**2.07**	**2.23**	**2.06**

Bold face fonts present the model that outperforms the best performing models of Table 2 and 3 for the 1-day and 10-days ahead, respectively.

Table 6: MCS *p*-values of the best performing models: Model-averaged forecasts (January, 2005 to July, 2013).

Model	Loss Function	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
		Implied Volatility Indices							
		One-day ahead							
SSA	MSE	0.0001	0.0001	0.2453*	0.0000	0.0002	0.0000	0.0001	0.0000
	MAE	0.0000	0.0000	0.0186	0.0000	0.0000	0.0000	0.0000	0.0000
SSA-HW	MSE	1.0000*	1.0000*	0.3604*	0.1360*	1.0000*	0.0720	0.0002	0.5968*
	MAE	0.8539*	1.0000*	0.0000	0.1944*	1.0000*	0.5014*	0.0000	0.3625*
ARI-SSA-HW	MSE	0.0016	0.5059*	1.0000*	1.0000*	0.7361*	1.0000*	1.0000*	1.0000*
	MAE	1.0000*	0.1249*	1.0000*	1.0000*	0.0723	1.0000*	1.0000*	1.0000*
		Ten-days ahead							
HW	MSE	0.0003	0.0028	0.7670*	0.0000	0.0002	0.0302	0.0000	0.0000
	MAE	0.0000	0.0000	0.1625*	0.0000	0.0000	0.0000	0.0000	0.0000
SSA-HW	MSE	0.0002	0.0171	0.3619*	0.0000	0.0020	0.0199	0.0000	0.0031
	MAE	0.0000	1.0000*	0.6613*	0.1713*	0.0787	0.0033	0.0024	0.0020
ARI-SSA-HW	MSE	1.0000*	1.0000*	1.0000*	1.0000*	1.0000*	1.0000*	1.0000*	1.0000*
	MAE	1.0000*	0.6324*	1.0000*	1.0000*	1.0000*	1.0000*	1.0000*	1.0000*

* denotes that the model belongs to the confidence set of the best performing models. The interpretation of the MCS *p*-value is analogous to that of a classical *p*-value; a $(1-a)$ confidence interval that contains the 'true' parameter with a probability no less than $(1-a)$.

Table 7: Direction of Change - One-day ahead (January, 2005 to July, 2013).

	Implied Volatility Indices							
Model	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
IV-HAR	0.5397	0.5270	0.5211	0.5336	0.5276	0.5315	0.5220	0.5202
IV-ARFIMA	0.5244	0.5347	0.5216	0.5364	0.5318	0.5315	0.5258	0.5164
HW	0.5077	0.4868	0.5053	0.5002	0.4995	0.5081	0.5158	0.5088
SSA	0.6789	0.6437	0.6207	0.6397	0.6336	0.6492	0.7204	0.6456
SSA-HW	0.7373	0.6992	0.6547	0.7044	0.6887	0.7169	0.7973	0.6922
I(1)	0.5785	0.4840	0.4646	0.4584	0.4577	0.4628	0.4618	0.4637
ARI(1,1)	0.5780	0.4926	0.4799	0.5296	0.4748	0.5243	0.4914	0.4907
FI(1)	0.5900	0.5318	0.5259	0.5450	0.5347	0.5372	0.5325	0.5287
ARFI(1,1)	0.5780	0.5122	0.5292	0.5093	0.5247	0.5148	0.5191	0.5059
ARI-IV-HAR	0.5431	0.5088	0.5005	0.5250	0.5157	0.5291	0.5105	0.5221
ARI-IV-ARFIMA	0.5258	0.5265	0.5115	0.5250	0.5166	0.5338	0.5096	0.5164
HAR-ARFIMA	0.5411	0.5328	0.5220	0.5393	0.5276	0.5372	0.5249	0.5164
ARI-SSA-HW	0.7340	0.6872	0.6379	0.6844	0.6811	0.6930	0.7677	0.6770

Table 8: Direction of Change - Ten-days ahead (January, 2005 to July, 2013).

	Implied Volatility Indices							
Model	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
IV-HAR	0.5630	0.5441	0.5598	0.5564	0.5779	0.5764	0.5488	0.5638
IV-ARFIMA	0.5749	0.5488	0.5655	0.5645	0.5703	0.5517	0.5360	0.5642
HW	0.6559	0.6498	0.6902	0.6545	0.6678	0.6300	0.6635	0.6406
SSA	0.4829	0.5005	0.5161	0.5308	0.5418	0.4720	0.4967	0.4917
SSA-HW	0.7180	0.7223	0.6917	0.8308	0.8783	0.6689	0.7739	0.7411
I(1)	0.4867	0.4512	0.4715	0.4564	0.4743	0.4568	0.4408	0.4661
ARI(1,1)	0.4905	0.4521	0.4682	0.4739	0.4796	0.4782	0.4673	0.4827
FI(1)	0.5820	0.5602	0.5740	0.5654	0.5827	0.5583	0.5445	0.5533
ARFI(1,1)	0.5815	0.5531	0.5802	0.5635	0.5822	0.5574	0.5427	0.5505
ARI-IV-HAR	0.5687	0.5275	0.5460	0.5488	0.5703	0.5697	0.5365	0.5614
ARI-IV-ARFIMA	0.5754	0.5531	0.5645	0.5602	0.5775	0.5398	0.5299	0.5505
HAR-ARFIMA	0.5763	0.5531	0.5669	0.5592	0.5798	0.5659	0.5398	0.5657
ARI-SSA-HW	0.7166	0.7133	0.6874	0.8265	0.8788	0.6618	0.7716	0.7378

Table 9: Naïve trading strategy results - One-day ahead (January, 2005 to July, 2013).

	Implied Volatility Indices															
Model	VCAC		VDAX		VFTSE		VIX		VSTOXX		VXD		VXJ		VXN	
IV-HAR	-0.0021		-0.0045		-0.0035		0.0000		-0.0039		-0.0012		-0.0033		-0.0030	
IV-ARFIMA	-0.0026		-0.0041		-0.0034		0.0001		-0.0029		-0.0009		-0.0037		-0.0032	
HW	-0.0065		-0.0079		-0.0068		-0.0059		-0.0076		-0.0059		-0.0060		-0.0053	
SSA	0.0213	***	0.0112	***	0.0113	***	0.0179	***	0.0128	***	0.0183	***	0.0225	***	0.0139	***
SSA-HW	0.0273	***	0.0167	***	0.0148	***	0.0249	***	0.0190	***	0.0255	***	0.0280	***	0.0193	***
I(1)	0.0016		-0.0067		-0.0056		-0.0053		-0.0066		-0.0054		-0.0055		-0.0076	
ARI(1,1)	0.0014		-0.0065		-0.0068		-0.0003		-0.0074		-0.0021		-0.0063		-0.0067	
FI(1)	0.0023		-0.0037		-0.0030		-0.0005		-0.0024		-0.0006		-0.0032		-0.0019	
ARFI(1,1)	0.0013		-0.0060		-0.0032		-0.0047		-0.0040		-0.0042		-0.0033		-0.0040	
ARI-IV-HAR	-0.0018		-0.0062		-0.0049		-0.0010		-0.0034		-0.0007		-0.0050		-0.0031	
ARI-IV-ARFIMA	-0.0027		-0.0046		-0.0029		-0.0009		-0.0037		-0.0006		-0.0053		-0.0033	
HAR-ARFIMA	-0.0013		-0.0045		-0.0032		0.0007		-0.0039		0.0000		-0.0029		-0.0038	
ARI-SSA-HW	0.0271	***	0.0155	***	0.0132	***	0.0225	***	0.0178	***	0.0227	***	0.0256	***	0.0179	***

Note: The numbers denote net average daily profits having deducted the transaction costs. *** denotes significance at 1% level.

Table 10: Naïve trading strategy results - Ten-days ahead (January, 2005 to July, 2013).

	Implied Volatility Indices									
Model	VCAC	VDAX	VFTSE	VIX		VSTOXX		VXD	VXJ	VXN
IV-HAR	-0.0005	-0.0014	-0.0009	-0.0012		-0.0011		-0.0007	-0.0012	-0.0014
IV-ARFIMA	-0.0005	-0.0012	-0.0008	-0.0009		-0.0014		-0.0012	-0.0016	-0.0013
HW	0.0019	0.0016	0.0028	0.0023		0.0024		0.0010	0.0021	0.0008
SSA	-0.0050	-0.0041	-0.0042	-0.0020		-0.0018		-0.0046	-0.0056	-0.0038
SSA-HW	0.0041	0.0027	0.0034	0.0070	***	0.0074	***	0.0024	0.0046	0.0039
I(1)	-0.0035	-0.0044	-0.0044	-0.0035		-0.0039		-0.0037	-0.0043	-0.0039
ARI(1,1)	-0.0035	-0.0043	-0.0047	-0.0033		-0.0037		-0.0036	-0.0037	-0.0037
FI(1)	-0.0002	-0.0011	-0.0004	-0.0009		-0.0006		-0.0009	-0.0015	-0.0012
ARFI(1,1)	-0.0003	-0.0011	-0.0001	-0.0010		-0.0006		-0.0010	-0.0013	-0.0012
ARI-IV-HAR	-0.0005	-0.0016	-0.0014	-0.0012		-0.0012		-0.0008	-0.0015	-0.0015
ARI-IV-ARFIMA	-0.0005	-0.0010	-0.0007	-0.0011		-0.0010		-0.0014	-0.0016	-0.0016
HAR-ARFIMA	-0.0004	-0.0012	-0.0007	-0.0011		-0.0010		-0.0009	-0.0014	-0.0013
ARI-SSA-HW	0.0041	0.0026	0.0033	0.0069	***	0.0074	***	0.0024	0.0046	0.0039

Note: The numbers denote net average daily profits having deducted the transaction costs. *** denotes significance at 1% level.

Table 11: Options straddles trading strategy results - One-day ahead (January, 2005 to July, 2013).

	Implied Volatility Indices							
Model	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
IV-HAR	0.3885	0.0394	0.3744	-0.2494	**0.6541**	-0.1577	0.1977	0.0153
IV-ARFIMA	0.4574	-0.0544	-0.2506	0.2125	0.1862	0.3394	0.2203	0.1313
HW	0.1623	0.0009	-0.6765	-0.0019	-1.0273	0.1923	-1.3628	0.1229
SSA	1.5345	0.5137	0.3608	1.8460	0.4259	1.5356	0.5840	**1.7529**
SSA-HW	**1.7328**	0.6248	**0.5820**	**1.8907**	0.3808	**1.8071**	0.7806	1.5875
I(1)	-2.0262	-0.7400	-0.3316	1.1024	-0.3824	0.8889	-0.2549	1.3423
ARI(1,1)	-1.8083	-0.4687	-0.2150	0.7909	-0.0032	0.7287	-0.7495	0.9944
FI(1)	-1.8142	0.1899	0.4554	-2.9625	0.0091	-2.7395	0.4828	-3.3940
ARFI(1,1)	-2.2580	-1.2324	-0.5838	-1.7980	-0.6758	-1.6637	**1.0054**	-1.6315
ARI-IV-HAR	0.6938	-0.7415	-0.4630	-0.7877	-0.2910	-0.7310	-0.3726	-0.7729
ARI-IV-ARFIMA	0.9511	-0.0024	-0.1147	-0.5308	-0.3434	-0.2983	-0.7126	-0.4818
HAR-ARFIMA	0.4385	0.7247	0.3523	-0.9026	0.4887	-0.9662	-0.7850	-0.9495
ARI-SSA-HW	1.5479	**1.1460**	0.5106	1.3904	0.5783	1.0644	0.9666	1.2831

Note: The numbers denote the cumulative returns, $\pi^{(j)} = \frac{1}{2}\sum_{t=1}^{\check{T}}\sum_{j^*=1}^{2}\pi_t^{(j,j^*)}$. Bold face fonts show the model with the highest profit levels.

Table 12: Options straddles trading strategy results - Ten-days ahead (January, 2005 to July, 2013).

	Implied Volatility Indices							
Model	VCAC	VDAX	VFTSE	VIX	VSTOXX	VXD	VXJ	VXN
IV-HAR	-0.5497	0.6566	0.8921	-0.5788	1.1247	0.3138	0.8479	-0.1423
IV-ARFIMA	-0.2667	0.2404	-0.0174	0.1151	0.4626	-0.2959	0.1056	-0.1485
HW	0.7565	-0.6068	-0.0908	1.3347	-0.1626	1.1698	-1.2018	0.9666
SSA	-2.0723	-0.3865	-0.7066	0.1139	0.2887	0.0829	-1.2985	-0.5859
SSA-HW	1.2334	0.1431	-0.1499	**2.9317**	-0.9638	**2.1831**	0.1474	**2.4445**
I(1)	0.3147	**0.9410**	1.0424	0.6917	1.0982	0.8261	-0.8306	0.9021
ARI(1,1)	0.2484	1.1948	**1.3231**	0.4784	**1.2422**	0.7619	-0.8407	0.5526
FI(1)	-1.9666	-1.5651	-1.6662	-3.0783	-1.6858	-2.6482	0.7632	-2.8646
ARFI(1,1)	-1.5126	-1.3459	-1.4235	-2.3952	-1.3117	-1.9734	**1.0257**	-2.0601
ARI-IV-HAR	0.9338	-0.2127	-0.0595	-0.5023	-0.2432	-0.5469	0.3673	-0.0450
ARI-IV-ARFIMA	0.3410	0.0971	0.4348	-0.6181	-0.0043	-0.2464	-0.0656	0.1175
HAR-ARFIMA	0.5185	0.5862	0.5435	-1.0072	0.2369	-0.6787	0.0039	-1.1432
ARI-SSA-HW	**2.0216**	0.2577	-0.1219	2.5146	-0.0819	1.0519	0.9763	2.0063

Note: The numbers denote the cumulative returns, $\pi^{(j)} = \frac{1}{2}\sum_{t=1}^{\check{T}}\sum_{j^*=1}^{2}\pi_t^{(j,j^*)}$. Bold face fonts show the model with the highest profit levels.

Appendix: Naive, ARFIMA and HAR models

(1) Naïve models

The I(1), ARI(1,1), FI(1) and ARFI(1,1) naïve models have been estimated in the following specifications.

a) The IV-I(1) model is estimated in the form:

$$(1 - L)(\log(IV_t)) = \beta_0 + \varepsilon_t, \tag{A.1}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and β_0 denotes parameter for estimation.

b) The IV-ARI(1,1) model is estimated in the form:

$$(1 - c_1 L)(\log(IV_t) - \beta_0) = \varepsilon_t, \tag{A.2}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and c_1 and β_0 denote parameters for estimation.

c) The IV-FI(1) model is estimated in the form:

$$(1 - L)^d(\log(IV_t) - \beta_0) = \varepsilon_t, \tag{A.3}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and β_0 and d denote parameters for estimation.

d) The IV-ARFI(1,1) model is estimated in the form:

$$(1 - c_1 L)(1 - L)^d(\log(IV_t) - \beta_0) = \varepsilon_t, \tag{A.4}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and c_1, β_0 and d denote parameters for estimation.

(2) Single HW and SSA models

The single HW model is similar to Eqs 4-7 of the main document, replacing $IV_{t,r}$ with IV_t; i.e. it is estimated for the implied volatility series rather than each one of the components of the implied volatility series.

For the SSA we follow the algorithm of Hassani and Thomakos (2010).

(3) IV-ARFIMA model

The long memory property of implied volatility indices makes the Autoregressive Fractionally Integrated Moving Average, or ARFIMA, model an appropriate framework for multiple-step-ahead implied volatility index, IV_t, predictions. The IV-ARFIMA(k,d,l) model for the discrete time t real-valued process $\log\left(IV_t\right)$ is utilized in the form[11]:

$$\begin{gathered}(1-C(L))(1-L)^d\left(\log\left(IV_t\right)-\boldsymbol{\beta}'\mathbf{x}_{t-1}\right)=(1+D(L))\varepsilon_t\,,\\ \varepsilon_t\sim N\left(0,\sigma_\varepsilon^2\right),\end{gathered}\tag{A.7}$$

where $\mathbf{x}_{t-1}=\left[1 \quad y_{t-1} \quad d_{t-1}y_{t-1}\right]'$ is the vector of explanatory variables, $\boldsymbol{\beta}$ is a vector of unknown parameters, and $C(L)=\sum_{i=1}^{k}c_iL^i$, $D(L)=\sum_{i=1}^{l}d_iL^i$ are polynomials with the parameters $c_1,...,c_k,d_1,...,d_l$ for estimation. The y_t denotes the log-returns of the underlying stock index and the d_t is a binary dummy variable, i.e. $d_t=1$, if $y_t>0$ and zero otherwise[12].

We define the orders of k and l of the IV-ARFIMA(k,d,l) model based on the Schwarz (1978) information criterion (for the total sample)[13], which is reported in Table A.1.

[TABLE A.1 HERE]

The IV-ARFIMA($2,d,1$) model is estimated for all the IV indices, except for the VCAC, VXN and VXJ, for which the IV-ARFIMA($2,d,2$) has been selected.

For the ARFIMA($2,d,1$) model the one-step-ahead logarithmic implied volatility, $\log\left(IV_{t+1|t}\right)$, is estimated as:

$$\log\left(IV_{t+1|t}\right)=\left(\hat{c}_1+\hat{c}_2L\right)\log\left(IV_t\right)+\left(1-\hat{c}_1L-\hat{c}_2L^2\right)\hat{\boldsymbol{\beta}}'\mathbf{x}_t+\sum_{j=1}^{\infty}A_jL^{j-1}\varepsilon_{t|t}+\sum_{j=0}^{\infty}A_jL^j\hat{d}_1\varepsilon_{t|t}\tag{A.8}$$

where $A_j\equiv\dfrac{\Gamma\left(j+\hat{d}\right)}{\Gamma\left(\hat{d}\right)\Gamma\left(j+1\right)}$, and $\varepsilon_{t|t}$ denotes the residual term at time t estimated based on the information set at time t, or $\varepsilon_{t|t}=\log\left(IV_t\right)-\left(\hat{c}_1+\hat{c}_2L\right)\log\left(IV_{t-1}\right)$

[11] The ARFIMA model was initially developed by Granger and Joyeux (1980).

[12] The dummy variable models the asymmetric relationship between volatility and lagged log-return; i.e. Degiannakis (2008b).

[13] The models were estimated in the ARFIMA package of Ox; see Doornik and Ooms (2006). The Schwarz information criterion (SBC) is computed from the Akaike information criterion (AIC) provided by ARFIMA package: $SBC=AIC+\breve{T}^{-1}q\left(\log\left(\breve{T}\right)-2\right)$, for $\breve{T}$ and q denoting the number of observations and parameters of the models (including the residuals' variance), respectively.

$-\left(1-\hat{c}_1 L-\hat{c}_2 L^2\right)\hat{\boldsymbol{\beta}}'\mathbf{x}_{t-1}-\sum_{j=0}^{\infty} A_j L^j \varepsilon_{t|t}-\sum_{j=0}^{\infty} A_j L^j \hat{d}_1 \varepsilon_{t-1|t}$.[14] The infinite expansion of the fractional differencing operator is approximated as (see Xekalaki and Degiannakis, 2010): $\sum_{j=0}^{\infty}\left(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}L^j\right)=1+\frac{1}{1!}dL+\frac{1}{2!}d(1+d)L^2-\ldots$. The parameters of the models $[\boldsymbol{\beta}, d, c_1,\ldots, c_k, d_1,\ldots, d_l]$ are re-estimated at each trading day.

The 10-step-ahead logarithmic implied volatility is estimated as[15]:

$$\log\left(IV_{t+10|t}\right)=\left(\hat{c}_1+\hat{c}_2 L\right)\log\left(IV_{t+9|t}\right)+\sum_{j=10}^{\infty} A_j L^{j-10}\varepsilon_{t|t}+\sum_{j=9}^{\infty} A_j L^{j-9}\hat{d}_1\varepsilon_{t|t}. \tag{A.9}$$

For $\varepsilon_t \sim N\left(0,\sigma^2_{t,\varepsilon}\right)$, the $\exp\left(\varepsilon_t\right)$ is log-normally distributed. Thus, Granger and Newbold (1976) showed that the exponential transformation of $\log\left(IV_{t+s|t}\right)$ is not theoretically optimal and proposed the $\exp\left(\log\left(IV_{t+s|t}\right)+\frac{1}{2}\hat{\sigma}^2_{t+s,\varepsilon}\right)$ as an optimal estimator of $IV_{t+s|t}$. However, Bårdsen and Lütkepohl (2011) provided both theoretical and empirical evidence that the optimal forecast will rarely result in RMSE reductions relative to the naïve forecast. They suggest that using the naïve forecast is the preferred option in applied work. For our study, we have proceeded to the estimation of both optimal and naïve forecasts. The adjustment factor $\frac{1}{2}\hat{\sigma}^2_{t+s,\varepsilon}$ is relatively too small, providing almost identical results. Hence, we conclude to estimate the s-trading-day-ahead implied volatility forecasts as:

$$IV_{t+s|t}=\exp\left(\log\left(IV_{t+s|t}\right)\right). \tag{A.10}$$

(4) IV-HAR model

The Heterogeneous Autoregressive, or HAR, model relates the current trading day's implied volatility with the daily, weekly and monthly implied volatilities. The autoregressive structure of the volatility over different interval sizes attempts to replicate the different perspectives that market participants may have on their

---

[14] Accordingly, the $\varepsilon_{t-1|t}$ denotes the residual term at time *t-1* estimated based on the information set at time *t*.

[15] The *s*-step-ahead forecast, for $s>2$, is $\log\left(IV_{t+s|t}\right)=\left(\hat{c}_1+\hat{c}_2 L\right)\log\left(IV_{t+s-1|t}\right)+\sum_{j=s}^{\infty} A_j L^{j-s}\varepsilon_{t|t}+\sum_{j=s-1}^{\infty} A_j L^{j-s+1}\hat{d}_1\varepsilon_{t|t}$.

investment horizon, which is the basic idea of the heterogenous market hypothesis in economic theory; see Müller *et al.* (1997).

The IV-HAR, model for the discrete time real-valued process $\log\left(IV_t\right)$ is defined as[16]:

$$\log\left(IV_t\right) =$$
$$w_0 + w_1 \log\left(IV_{t-1}\right) + w_2\left(5^{-1}\sum_{j=1}^{5}\log\left(IV_{t-j}\right)\right) + w_3\left(22^{-1}\sum_{j=1}^{22}\log\left(IV_{t-j}\right)\right) + \varepsilon_t, \qquad \text{(A.11)}$$
$$\varepsilon_t \sim N\left(0,\sigma_\varepsilon^2\right),$$

where the w_0, w_1, w_2, w_3 are the unknown parameters to be estimated[17].

The IV-HAR model forecast for the 1-day-ahead is computed as:

$$IV_{t+1|t} =$$
$$\exp\left(\hat{w}_0 + \hat{w}_1\log\left(IV_t\right) + \hat{w}_2\left(5^{-1}\sum_{j=1}^{5}\log\left(IV_{t-j+1}\right)\right) + \hat{w}_3\left(22^{-1}\sum_{j=1}^{22}\log\left(IV_{t-j+1}\right)\right) + \tfrac{1}{2}\hat{\sigma}_{t,\varepsilon}^2\right) \qquad \text{(A.12)}$$

The 10-days-ahead logarithmic implied volatility, based on IV-HAR model, is computed as:

$$\log\left(IV_{t+10|t}\right) = \hat{w}_0 + \hat{w}_1\log\left(IV_{t+9|t}\right) + \hat{w}_2\left(5^{-1}\sum_{j=1}^{5}\log\left(IV_{t-j+10|t}\right)\right) +$$
$$\hat{w}_3\left(22^{-1}\left(\sum_{j=1}^{9}\log\left(IV_{t-j+10|t}\right) + \sum_{j=10}^{22}\log\left(IV_{t-j+10}\right)\right)\right). \qquad \text{(A.13)}$$

Table A.1: The SBC criterion for various orders of the IV-ARFIMA(*k,d,l*) model.

	k=0 *l*=0	*k*=0 *l*=1	*k*=1 *l*=0	*k*=1 *l*=1	*k*=2 *l*=1	*k*=1 *l*=2	*k*=2 *l*=2	*k*=3 *l*=2	*k*=2 *l*=3
VIX	-2.338	-2.528	-2.607	-2.650	**-2.664**	-2.656	-2.661	-2.659	-2.659
VSTOXX	-2.415	-2.683	-2.817	-2.844	**-2.863**	-2.853	-2.861	-2.858	-2.858
VFTSE	-2.292	-2.549	-2.690	-2.724	**-2.739**	-2.730	-2.735	-2.733	-2.735
VDAX	-2.609	-2.906	-3.077	-3.108	**-3.130**	-3.114	-3.127	-3.125	-3.125
VCAC	-2.400	-2.609	-2.714	-2.759	-2.763	-2.760	**-2.766**	-2.758	-2.764
VXN	-2.606	-2.848	-2.966	-3.006	-3.018	-3.011	**-3.019**	-3.013	-3.016
VXD	-2.372	-2.564	-2.650	-2.698	**-2.712**	-2.702	-2.709	-2.706	-2.706
VXJ	-2.353	-2.483	-2.547	-2.618	-2.622	-2.620	**-2.622**	-2.619	-2.620

Bold face fonts present the best order of the IV-ARFIMA(*k,d,l*) model.

[16] The HAR model initially developed by Corsi (2009).

[17] The HAR model could be extended to accommodate heteroscedasticity in the error term, as in Corsi *et al.* (2005). However, the modeling of volatility of realized volatility is out of the scope of the paper.