Erasmus University Rotterdam

Erasmus School of Economics

Master Thesis Quantitative Finance

The statistical and economic relevance of out-of-sample forecasts of implied volatility surfaces of equity options

Author: E.N. van Bruggen (400380)

Supervisor: dr. M. Grith

Second assessor: dr. M. van der Wel

September 16, 2018

Abstract

The implied volatility surface (IVS) is a mapping of implied volatilities of options as a function of the moneyness and time-to-maturity. Capturing and forecasting the dynamics of these surface can contribute to trading and hedging strategies, as it contains information about the expected market volatility. I propose both a two-step approach based on principal components analysis (PCA) and a one-step approach based on a state-space model. Exogenous variables are included at a later stage to improve the predictions. The most relevant factors and exogenous variables are selected by the least angle regressions (LARS) algorithm. There are three main findings. First, both of the approaches capture the predictability of the IVS of equity options in the information technology (IT) sector. The one-step approach is the best performing averaged over time for each individual equity option in terms of performance evaluation measures. Second, the consideration of exogenous variables improves the predictions of both approaches in terms of statistical performance. The expected inflation and interest rates are found to contribute the most to the prediction of the IVS of equity options. Third, when disregarding transaction costs, the one-step ahead forecasts of both approaches are able to produce positive returns for an ATM straddle trading strategy. However, when accounting for transaction costs, these potential profits disappear.

Keywords: Implied Volatility Surface (IVS), Factor models, PCA, Kalman filter, LARS algorithm, ATM straddle strategy

1 Introduction

Implied volatility is a measure of the volatility that is obtained from observed market prices of assets. The Black–Scholes implied volatility of a call option is defined as the value of the volatility parameter such that the market price equals the price given by the Black–Scholes model for options pricing. In the Black-Scholes model, the volatility is constant with time. However, in reality, there are two key dimensions on which volatilities implied by options written on the same underlying asset depend: moneyness and time-to-maturity. Mapping the implied volatility as a function of these two variables generates an implied volatility surface (IVS). Changes in these variables cause the IVS to shift over time. Shifts in implied volatility are correlated across strikes and maturities, causing the IVS to be relatively smooth. This correlation is due to the fact that option prices for different moneyness and maturities depend on the same underlying stock. As implied volatility is a transformation of the option prices, this co-movement will also be present in the IVS. These characteristics suggest that the dynamics of the IVS are driven by a small number of factors and make parsimonious modeling possible.

There exists a clear relevance for market participants to produce reliable IVS predictions, as it provides an up-to-date indication of the expected market volatility [Bernales and Guidolin, 2014]. Risk managers use implied volatility as a measure of risk in the financial market, whereas in option pricing it characterises the future beliefs of market participants. Therefore accurate forecasts can contribute to trading and hedging strategies. The attempt of this thesis is to further develop forecasting abilities of dynamic factor models where the latent factors drive the dynamics of the surface. The statistical relevance of including exogenous variables for the forecasts of equity options will be established. It is then investigated whether the out-of-sample forecasts have economic value when applied to an ATM straddle trading strategy.

Most studies about the predictability of the IVS focuses on index options on the S&P 500, since these are actively traded and the options are available for many strike prices and maturities. Options on stocks are less liquid than index options, therefore results found in these papers mostly do not hold for equity options. This is due to the slow updating process of the equity option IVS caused by the low trading frequency. Because indexes are composed of several equities, positions taken in an index are typically motivated by macroeconomic expectations or hedging purposes, rather than microeconomic analysis of the underlying stocks. Predicting the IVS of individual equity options can thus offer insights more widely applicable to trading strategies. The volatility of the equities asset class is difficult to charaterise, since the value of different equities may be driven by different factors. Within a sector similar equity options IVS dynamics are expected, since the stocks typically share value drivers. To gain a better understanding of the relevance of forecasts of the IVS of equity option this paper focuses on equity options within one sector, the IT sector. Some of the world's largest companies by market capitalisation are IT companies. These stocks are characterised by high trading volumes. Stocks in the IT sector also have a relatively high average volatility. Investors in this sector typically expect rapid growth and speculate on the profitability of digital innovations.

Various models of surface construction and its dynamics have been developed and improved over time. A number of papers have explored the index IVS movements using principal component analysis (PCA) [Cont et al., 2002, Skiadopoulos et al., 2000] or using a Kalman filter approach [Bedendo and Hodges, 2009, van der Wel et al., 2016]. Moreover, there is empirical evidence of the predictability of the IVS [Dumas et al., 1998, Cont et al., 2002, Goncalves and Guidolin, 2006, Fengler et al., 2007]. These papers focus mainly on the predictability of the IVS of index options, such as S&P 500 index options. Studies that explore the predictability of the IVS of equity options are sparse. Goncalves and Guidolin [2006] forecast the IVS of index options using a two-step framework. First, they model the IVS of the S&P 500 index options by fitting on a daily basis a parametric model that describe the implied volatilities as a function of moneyness and time-to-maturity. In the second step, they fit a VAR model to produce the forecasts in which the factors are the time series coefficients estimated in the first step. They found that the performance of this model is superior to hard-to-beat benchmarks, such as a random walk model. The existence of dynamic relationships between the IVS of equity options and the IVS of index options has been investigated by Bernales and Guidolin [2014] using the same two-step approach as Goncalves and Guidolin [2006]. They drew two key conclusions. First, the IVS of equity options can be forecasted. Second, the movements in the IVS of index options provide useful information to forecast the IVS of equity options. An out-of-sample forecasting approach based on the factors estimated with PCA is taken by Chalamandaris and Tsekrekos [2010], who studied IVS dynamics of 25 over-the-counter currency options. In the three aforementioned studies, the trading profits created with the forecasts of the IVS are offset by transaction costs. Therefore, it can not yet be concluded that the predictable movements support portfolio decisions. This is a motivation to investigate whether alternative approaches of forecasting the IVS of equity options may not only hold as a statistical fact but also be relevant from an economic perspective.

Recently, van der Wel et al. [2016] provided an approach to estimate the factors and their dynamics in one step by means of the Kalman filter. This one-step approach estimates the factors and the IVS simultaneously so that all the uncertainties of the factor forecasts are taken into account. The focus of the study of van der Wel et al. [2016] is the construction of the IVS based on the Kalman filter; they are not concerned with whether the dynamic factor models accurately forecast the IVS. Bedendo and Hodges [2009] produced returns for options portfolios based on forecasts of the volatility skew using a model based on the Kalman filter. They do not consider the volatility skew across various maturities.

Apart from constructing and forecasting the IVS by means of factors models, several papers [Franks and Schwartz, 1991, Mixon, 2002] related the changes in implied volatility surfaces to observable state variables. This is relevant since asset prices react to economic news, which suggests the presence of observable influences. Chen et al. [1986] have identified variables which influence stock market returns. Mixon [2002] found evidence that observable state variables can to some extent explain the movements in the IVS of index options, with the most predictive power found in index returns.

In this study, I will present two approaches to create forecasts of the IVS of equity options. In the first approach I will consider a two-step framework. First, the IVS is fitted by a factor model. Then, the IVS is predicted based on the forecasts of the factors. The set of factors is constructed by a principal component analysis as in Chalamandaris and Tsekrekos [2010]. After the IVS is fitted, the time-variation of the IVS is captured by the factors that are assumed to follow a VAR model. I will then extend the VAR model with factors that characterise the IVS of index options and exogenous variables. Motivated by Chen et al. [1986], Franks and Schwartz [1991], Bernales and Guidolin [2014] and Mixon [2002] the following exogenous variables are chosen: the contemporaneous and lagged index return of the S&P 500, real interest rates, inflation, corporate/government bond yield spread, and factors that characterise the IVS of index option. The addition of the exogenous variables will be evaluated on a daily basis. A common approach is to use monthly inflation data; I instead use the daily expected inflation, $E[\pi_t]$. This represents what market participants expect the inflation to be over the next 5 years, on average. This variable has not been used for IVS predictions in previous studies.

Several papers [Bair et al., 2006, Bai and Ng, 2008, Stock and Watson, 2012, Kelly and Pruitt, 2015] stress that statistical factor models construct the factors regardless of the observed time-series to be predicted. They found that the selection of more relevant factors improves the forecast performance of the factor models. Therefore I use the Least Angle Regression (LARS) algorithm introduced by Efron et al. [2004] to select the set of most relevant factors for predicting the implied volatility. The factors are selected from the factors created by PCA and the exogenous variables in the two-step approach. The LARS algorithm is efficient in terms of computation time compared to other selection methods. Furthermore, LARS weights all the variables according to their relevance, rather than leaving some variables out completely.

In the second approach, I will consider a state-space model for the implied volatility inspired by van der Wel et al. [2016]. This is referred to as the one-step approach since it estimates the factors and forecasts the IVS of equity options simultaneously by use of the Kalman filter. The Kalman filter was introduced by Kalman [1960]. It is a recursive procedure that provides an optimal forecast of the factors one step ahead given the currently available information. In this way, the filter makes the best use of the available information in time sequence for the prediction of the factors and of the implied volatility. Similar to the two-step approach I extend the state-space model with exogenous variables. There is no algorithm to select state-space factors, since these are typically unobserved. Therefore the LARS algorithm cannot be used to select the exogenous variables along the state space factors. The exogenous variables selected in the two-step approach, are also used to extend the state-space model in the one-step approach.

For comparative purposes, I will also estimate two benchmark models. The first benchmark model is based on the paper of Goncalves and Guidolin [2006] who use a parametric model to fit the IVS. This model is used by a number of studies, among which Bernales and Guidolin [2014], and shown to be superior to the random walk model. The factors are related to strike prices and time-to-maturities since these are two key dimensions for option prices. In this paper is concentrated on short-horizon forecasts up to five days ahead. Related studies found that for short-horizon forecast the random walk model is hard to beat. Therefore it is also relevant to include the naive random walk model as a second benchmark model. The forecast performance of the approaches will be evaluated based on three performance criteria. The economic value of equity IVS forecasts from the two approaches is assessed by executing a simple ATM straddle trading strategy.

For the empirical study, I use a dataset of daily implied volatilities for options on equity. The dataset includes options on the ten stocks with the largest market capitalisation in the S&P 500 in the information technology sector from September 1, 2004, through December 31, 2017. All considered companies have continuous listing during the sample period and are actively traded. Additionally, I have data on the daily implied volatilities for European options on the S&P500 index. I have daily data of interest rates, expected inflation, index returns (S&P500), and corporate/government bond yield spread to analyse the possible relationship between the IVS of equity options and these exogenous variables.

The results confirm that the proposed forecast approaches and the selected exogenous variables can forecast the IVS of equity options. This conclusion is supported by three main findings in this thesis. First, both of the approaches capture the predictability of the IVS of equity options in the IT sector. The one-step approach is the best performing averaged over time for each individual equity option in terms of performance evaluation measures. Second, the consideration of exogenous variables improves the predictions of both approaches in terms of statistical performance. The expected inflation and interest rates are found to contribute the most to the prediction of the IVS of equity options. Third, when disregarding taking transaction costs, the one-step ahead forecasts are able to produce positive returns for an ATM straddle trading strategy. However, when accounting for transaction costs, these potential profits disappear.

This study contributes to the existing literature in three ways. First, this is one of the first papers to use a state space model to forecast the IVS. The application to out-of-sample forecasting for IVS of equity options of companies operating in the same sector are relatively rare. Moreover, research on the interactions between the IVS of equity options and exogenous variables is sparse.

This paper is structured as follows: section 2 describes the different approaches to predict the implied volatility surface. Section 3 elaborates on the data used for the empirical application. Section 4 will describe the empirical results of the approaches. Section 5 concludes and discusses potential subjects for further research.

2 Methodology

This section describes two approaches to make out-of-sample predictions of the IVS of equity options. In the approaches, I use as observation vectors, T time series of N units of log implied volatility. Each unit refers to an option on a stock with a different moneyness (m) and time-to-maturity (τ). The grid of observed implied volatilities are both irregular and changing with time as the level of the underlying fluctuates [Cont et al., 2002]. To deal with this unbalanced panel I follow the approach of van der Wel et al. [2016]. I create several maturity-moneyness groups. Within each maturity-moneyness group, I select the option that has the smallest Euclidean distance to the midpoint of both dimensions to obtain a balanced panel of similar options across time. If there is no options data within a group on a specific day, these missing values are generated by selecting a neighbor defined as the option with the smallest Euclidean distance to both midpoints across all other groups at that day [van der Wel et al., 2016]. This may cause some extra bias in the predictions. In this paper I use the following notation, $\boldsymbol{IV}^k$ is the $(N \times T)$ matrix of log implied volatilities. Let $\boldsymbol{IV}_t^k$ is a column vector denoting all N log implied volatilities for options on stock k at time t, while $\boldsymbol{IV}_i^k$ is a row vector denoting the log implied volatilities over all days T for option i on stock k. Specifically, I select three maturity groups ($\tau = 3$) classified as short-term, medium-term, and long-term. Next, I select six moneyness groups ($M = 6$) of at-the-money, out-of-the-money, and deep out-of-the-money call and put options, which results in $N = 18$ different maturity-moneyness groups. Time-to-maturity (τ) is measured in days. Following the approaches of Bollen and Whaley [2004] and van der Wel et al. [2016] moneyness is defined in terms of Δ. The Δ of an option is the sensitivity of the option price with respect to movements in the price of the underlying. More common is to define moneyness as the ratio of the strike price to the forward price. However, the advantage of using Δ is that it considers the volatility of the underlying asset and the remaining time-to-maturity.

$$\boldsymbol{IV}_t^k = \begin{pmatrix} \log(IV^k_{\tau_1,m_1,t}) \\ \vdots \\ \log(IV^k_{\tau_\tau,m_1,t}) \\ \log(IV^k_{\tau_1,m_2,t}) \\ \vdots \\ \log(IV^k_{\tau_\tau,m_M,t}) \end{pmatrix} \tag{1}$$

I am interested in the h-step ahead out-of-sample forecast of the implied volatility, $\hat{\boldsymbol{IV}}^k_{t+h|t}$. Therefore the logarithm of the implied volatility is used as dependent variable since this has the advantage of always producing non-negative forecasts. Finally, I convert the forecasts to standard implied volatilities by using log normality property $\hat{\boldsymbol{IV}}^k_{t+h|t} = \exp\Big(\log(\hat{\boldsymbol{IV}}^k_{t+h|t}) + \frac{1}{2}\mathrm{Var}_t(\log(\hat{\boldsymbol{IV}}^k_{t+h|t}))\Big)$.

In subsection 2.1 I discuss the two-step approach, where first the IVS is fitted and subsequently is predicted. The one-step approach that fits and predicts the IVS simultaneously is discussed in subsection 2.2. Subsection

2.3 describes the benchmark models, followed by the evaluation measures in subsection 2.4. Finally, section 2.5 describes the trading strategy that applies the predictions of the IVS.

2.1 Two-step approach

In the two-step approach, I first fit the IVS based on factors obtained from a principal component analysis. The time-variation of the IVS is captured by the factors following a VAR model. Then the predictions of the factors are used to predict the IVS h-step ahead. The VAR factor model is then extended by adding exogenous variables. I use the LARS algorithm to select the most relevant factors for the prediction of the implied volatility from the factors created by PCA and the exogenous variables. Then, I use the first rolling window as an estimation period for the factors. I make predictions for the rest of the dataset using a rolling window. The rolling window has a length of 200 days. The model is reestimated for every day. Log implied volatilities can be modeled using first differences or levels. To select the right modeling approach, unit root tests will be performed on the univariate factors and the exogenous variables. The augmented Dickey-Fuller (ADF) test will be performed. The ADF test tests the null hypothesis that a unit root is present in a time series against the alternative hypothesis of a stationary time series. The test statistic is a negative number. The more negative the statistic, the stronger the rejection for the presence of unit roots. In addition, the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test is performed. The KPSS has the presence of unit roots as an alternative hypothesis. A failure to reject the null-hypothesis indicates stationarity.

2.1.1 Fitting IVS

I consider a factor representation to fit the daily implied volatility surfaces. The set of factors are created by a principal component analysis.

$$\boldsymbol{IV}_t^k = \boldsymbol{\theta}_1^k F_{1,t}^k + \cdots + \boldsymbol{\theta}_r^k F_{r,t}^k + \boldsymbol{\epsilon}_t = \boldsymbol{\Theta}^k \boldsymbol{F}_t^k + \boldsymbol{\epsilon}_t \tag{2}$$

Let $\boldsymbol{F}_t^k = (F_{1,t}^k, F_{2,t}^k, ..., F_{r,t}^k)'$ be r factors of stock k created by PCA, and let $\boldsymbol{\Theta}^k = (\boldsymbol{\theta}_1^k, \boldsymbol{\theta}_2^k, ..., \boldsymbol{\theta}_r^k)'$ be the $(N \times r)$ matrix of parameters for stock k. The factors $\boldsymbol{F}_t^k$ are time variant. The factor loadings $\boldsymbol{\Theta}^k = (\boldsymbol{\theta}_1^k, \boldsymbol{\theta}_2^k, ..., \boldsymbol{\theta}_r^k)'$ are fixed over time.

The aim of a principal component analysis is to reduce the data dimension to $r < N$ in such a way that the r factors capture the bulk of the variation in the dataset. The factors are created by identifying the linear combinations of the data that are uncorrelated and have maximum variance. This involves solving the following non-linear least square problem [Stock and Watson, 2006]:

$$\min_{\boldsymbol{F}_1^k, \cdots, \boldsymbol{F}_T^k, \boldsymbol{\Theta}^k} \quad T^{-1} \sum_{t=1}^{T} (\boldsymbol{IV}_t^k - \boldsymbol{\Theta}^k \boldsymbol{F}_t^k)'(\boldsymbol{IV}_t^k - \boldsymbol{\Theta}^k \boldsymbol{F}_t^k) \quad \text{subject to} \quad \boldsymbol{\Theta}^{k'} \boldsymbol{\Theta}^k = \boldsymbol{I}_r \tag{3}$$

The solution of this optimisation is given by $(\hat{\boldsymbol{\Theta}}^k, \hat{\boldsymbol{F}}_t^k)$ where $\hat{\boldsymbol{\Theta}}^k$ are r eigenvectors corresponding to the eigenvalues of the covariance matrix of $\boldsymbol{IV}_t^k$ and the factors are estimated by $\hat{\boldsymbol{F}}_t^k = \hat{\boldsymbol{\Theta}}^{k'} \boldsymbol{IV}_t^k$. A general explanation of PCA and the proof of this result is found in appendix A.1.1.

The common approach of a PCA is to select the r factors with the largest eigenvalues. However, the factors with the largest eigenvalues are not necessarily the best predictors for the variable to be predicted. Therefore I select the most relevant factors for the prediction of the IVS with the LARS algorithm of Efron et al. [2004]. The selection algorithm is discussed in the subsection 2.1.3. The number of factors r that is selected will be empirically determined for each rolling window with a maximum of $r = 5$ factors. The number of factors that give the best prediction based on in-sample fit is selected.

2.1.2 Modeling the time-variation of the IVS

The time-variation in the IVS is captured by the dynamics of the factors identified. To model the dynamics of the estimated factors I assume they follow a VAR process with p lags of the factors. A VAR model has proven to be a reasonable approach to model the predictability in the IVS [Chalamandaris and Tsekrekos, 2009].

$$\hat{\boldsymbol{F}}_t^k = \boldsymbol{c}^k + \sum_{j=1}^{p} \boldsymbol{\Phi}_j^k \hat{\boldsymbol{F}}_{t-j}^k + \boldsymbol{v}_t \qquad \boldsymbol{v}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}^k_{\hat{\boldsymbol{F}}^k \hat{\boldsymbol{F}}^k}) \tag{4}$$

where $\hat{\boldsymbol{F}}_t^k$ is a $(r \times 1)$ vector with factors obtained via the optimisation in equation 3. To ensure the right modeling approach for the factors, unit root tests will be performed on the univariate factors. Equation 4 is estimated with OLS. The appropriate number of p lags is selected based on the Bayesian information criterion (BIC) for every rolling window.

Motivated by Mixon [2002] and Bernales and Guidolin [2014] I include exogenous variables to improve the forecasts of the factors, so that equation 4 becomes a VARX model. The exogenous variables considered are:

- The contemporaneous and lagged index returns of the S&P 500 options, respectively $\boldsymbol{r}_t^{SPX}$ and $\boldsymbol{r}_{t-1}^{SPX}$
- The real interest rate in the USA, $\boldsymbol{I}_t$
- The expected inflation rate in the USA, $\boldsymbol{\pi}_t$
- The corporate/government bond yield spread, $\boldsymbol{Y}_t$
- The factors created by PCA that characterise the IVS of S&P 500 index options, $\boldsymbol{F}_t^{SPX}$

The definition and source of the exogenous variables is found in table 7 in appendix A.2. The model introduced below in equation 5 is the VARX model, which is used to forecast the factors that drive the IVS of equity options using the additional exogenous variables.

$$\hat{\boldsymbol{F}}_t^k = \boldsymbol{c}^k + \sum_{j=1}^{p} \boldsymbol{\Phi}_j^k \hat{\boldsymbol{F}}_{t-j}^k + \boldsymbol{\gamma}^k \boldsymbol{X}_t + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \boldsymbol{\Omega}) \tag{5}$$

where $\hat{\boldsymbol{F}}_{t-j}^k$ represents the factors created by a PCA to the covariance matrix of individual equity option k and its p lags. $\boldsymbol{X}_t = [\boldsymbol{r}_t^{SPX}, \boldsymbol{r}_{t-1}^{SPX}, \boldsymbol{I}_t, \boldsymbol{\pi}_t, \boldsymbol{Y}_t, \boldsymbol{F}_t^{SPX}]$ is the set of exogenous variables. To assess the relevance of the exogenous variables for predicting the IVS of equity options, they are considered alongside the factors created by PCA. The aggregate of these two sets, $[\hat{\boldsymbol{F}}_t^k, \boldsymbol{X}_t]$, is used as input for the LARS algorithm. The five that are found to be most relevant are included. The selection procedure is repeated for every rolling window. The factors and exogenous variables will be tested for unit roots to determine whether it is necessary to take the first differences of the log implied volatilities to ensure stationarity. Equation 5 is estimated with OLS. I select the number of lags p for the factors via the minimisation of the Bayes–Schwarz criterion for every rolling window.

2.1.3 Selecting factors

As explained in the introduction and subsection 2.1.1, the factors obtained by a PCA are created irrespectively of the variable to be predicted. Therefore I select the most relevant factors for the prediction of the IVS by the Least Angle Regressions (LARS) algorithm introduced by Efron et al. [2004]. When exogenous variables are considered for the prediction of the factors, the factors created by PCA along with the exogenous variables are used as input for the LARS algorithm. The LARS algorithm models the coefficients of the included variables jointly, in order to prevent that valuable information, is lost. The LARS coefficients are estimated for every rolling window, to obtain the most relevant set of predictors. This ensures that the factors/exogenous variables used for the prediction of the IVS can change over time.

The intuition behind the LARS algorithm is as follows: it starts off with an empty set. At each step, it adds the variable with the smallest angle, which equates to the variable with the largest correlation. From this variable, it increases the coefficient until there is another variable with the same correlation. The algorithm proceeds until all the variables are added according to their correlation with the variable to be predicted. Figure 2 from the paper of Efron et al. [2004] gives a simple visualisation of the algorithm with two variables. It starts with all coefficients equal to zero and finds the most correlated predictor, x_1. The algorithm takes the largest step possible in the direction of this predictor μ_1 until some other predictor, x_2, has as much correlation. At this point, it changes in the direction of u_2, which is equiangular between the two predictors.

The following paragraphs give a formal description of the LARS algorithm, customised for the application in this study. For a general explanation of the LARS algorithm, similar to Efron et al. [2004], I refer to appendix A.1.2. I define $\hat{\boldsymbol{IV}}_{q,t}^k$ as the current estimate of the implied volatility with q factors at time t. The

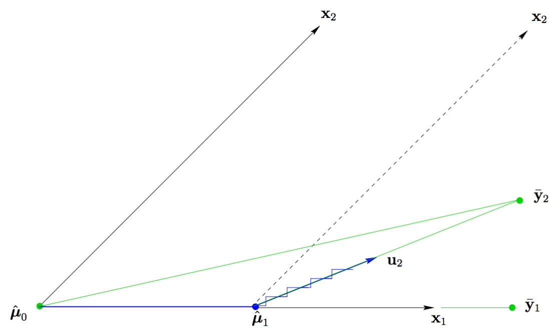


Figure 2: Visualisation of Least Angle regressions

corresponding current correlation is defined by $\hat{c}_t^k = \boldsymbol{F}_t^{k'}(\boldsymbol{IV}_t^k - \boldsymbol{I\hat{V}}_{q,t}^k)$, where $\boldsymbol{F}_t^k$ is the $(T \times q)$ matrix of factors. The time-subscript t is included here since the LARS algorithm is performed for every rolling window to find the most relevant factors. Since the same algorithm is applied to each rolling window, the time subscript t is left out for the remainder of the description of the algorithm. In order for the algorithm to work well, the input variables need to be standardised and the variable of interest need to be demeaned. The algorithm starts with $\boldsymbol{IV}_0^k = 0$. For all the factors $i = 1, \cdots n$ the correlation between $\boldsymbol{F}_i^k$ and $\boldsymbol{IV}_q^k$ can be calculated as the inner product since the factors are standardised and the implied volatility is demeaned. This implies that all means are zero and the standard deviations equal to 1.

$$\begin{aligned} \boldsymbol{c}_j^k &= \text{corr}(\boldsymbol{F}_j^k, \boldsymbol{IV}_q^k) = \frac{\text{cov}(\boldsymbol{F}_j^k, \boldsymbol{IV}_q^k)}{\sigma_{F_j^k}\sigma_{IV_q^k}} = \frac{\text{E}[(\boldsymbol{F}_j^k - \text{E}(\boldsymbol{F}_j^k))(\boldsymbol{IV}_q^k - \text{E}(\boldsymbol{IV}_q^k))]}{\sigma_{F_j^k}\sigma_{IV_q^k}} \\ &= \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{F}_{ij}^k - \hat{\boldsymbol{F}}_j^k)(\boldsymbol{IV}_{i,q}^k - \hat{\boldsymbol{IV}}_q^k) = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{F}_{ij}^k\boldsymbol{IV}_{i,q}^k \end{aligned} \tag{6}$$

First, the algorithm selects the factors that have the highest absolute correlations with the current residual $\hat{c} = \boldsymbol{F}^{k'}(\boldsymbol{IV} - \boldsymbol{I\hat{V}}_q)$. These are then added to the active set A.

$$A = [j : |\hat{c}_j| = \hat{C}] \qquad \hat{C} = \max_j |\hat{c}_j| \tag{7}$$

where the index $j = 1, \cdots, n$ refers to the n factors. When the factor is added, the algorithm has to decide in which direction to proceed. Therefore the algorithm tends to find the vector which has equal angles with the factors in the active set, also referred to as the equiangular vector.

$$\boldsymbol{u}_A = \boldsymbol{F}_A\boldsymbol{w}_A \qquad \text{where } \boldsymbol{w}_A = (\boldsymbol{\iota}_A'(\boldsymbol{F}_A'\boldsymbol{F}_A)_A^{-1}\boldsymbol{\iota}_A)^{-1/2}(\boldsymbol{F}_A'\boldsymbol{F}_A)^{-1}\boldsymbol{\iota}_A \tag{8}$$

$\boldsymbol{u}_A$ is a linear combination of the columns of $\boldsymbol{F}_A$ scaled by $\boldsymbol{w}_A$. $\boldsymbol{w}_A$ determines the direction of the prediction. $\boldsymbol{\iota}_A$ represents a column vector of ones of length $|A|$. All angles between the vectors are assumed to be less than 90 degrees, meaning that all correlations are positive. The next step of the LARS algorithm is to determine how far γ_A the algorithm has to proceed along $\boldsymbol{u}_A$ for the new prediction.

$$\boldsymbol{I\hat{V}}_{q+1} = \boldsymbol{I\hat{V}}_q + \hat{\gamma}_A\boldsymbol{u}_A \tag{9}$$

To determine $\hat{\gamma}_A$ the purpose of the algorithm needs to be considered. The algorithm needs to add one extra variable to the set A only if there is a variable that has as much correlation as the ones from the active set. In other words, $\boldsymbol{\gamma}_A$ needs to be chosen such that there exists a variable with $|c_j| = |c_i|$, where i is element of the active set and j is not in the active set. When a factor $\boldsymbol{F}_i$ is in the active set A the current correlations is $|c_i| = \pm C - \gamma_A w_A$. When a factor $\boldsymbol{F}_i$ is not in A the current correlations are $|c_j| = |\hat{c}_j - \gamma_A a_j|$. Since γ_A always has to be greater than zero the solution to the equality becomes:

$$\hat{\gamma} = \min_{j \in A^C}{}^{+}\Big(\frac{\hat{C} - \hat{c}_j}{\boldsymbol{H}_A - a_j}, \frac{\hat{C} + \hat{c}_j}{\boldsymbol{H}_A + a_j}\Big) \tag{10}$$

where $\boldsymbol{H}_A = (\boldsymbol{\iota}'_A(\boldsymbol{F}'_A\boldsymbol{F}_A)^{-1}_A\boldsymbol{\iota}_A)^{-1/2}$. Furthermore, $\min^+$ implies that the minimum is taken only over the components that are positive, within every j. So $\hat{\gamma}$ is the smallest value such that there is a factor $\boldsymbol{F}_j$, not yet in the active set, that has the same correlation with the residual.

2.1.4 Predicting the IVS

I make h-step ahead predictions of the IVS of equity options using the factor decomposition in equation 2 and the models of the factors dynamics in equation 4 and equation 5. The predictions will be mainly for a short-horizon, $h = [1, 3, 5]$ days, since previous studies found that the forecast accuracy diminishes when forecast horizon increases [Chalamandaris and Tsekrekos, 2010, Bernales and Guidolin, 2014]. To form a feasible h-days ahead forecast of the implied volatility surface of options on stock k, I need forecasts of the factors $\hat{\boldsymbol{F}}^k_{t+h|t}$. I use the first rolling window of daily surfaces of the dataset to extract the factors $\hat{\boldsymbol{F}}^k_t$ and loadings $\hat{\boldsymbol{\Theta}}^k = [\hat{\theta}^k_1, \cdots, \hat{\theta}^k_N]$ as in equation 2. I obtain the h-step ahead forecast directly from a long-horizon forecasting equation instead of a sequence of one-step ahead forecasts. The following approach based on Chalamandaris and Tsekrekos [2010], is applied to create h-step ahead forecasts of the IVS:

Let $\hat{\boldsymbol{\beta}}^k = [\hat{c}^k, \hat{\boldsymbol{\Phi}}^k_j]$ be the coefficients of the projection of $\hat{\boldsymbol{F}}^k_t$ on $\hat{\boldsymbol{F}}^k_{t-h}$ and its $\hat{p}$ lags as in equation 2. Obviously for a one-day ahead forecast, $h = 1$, the coefficients include $\hat{c}^k$ and $\hat{\Phi}^k_j$ with $j = 1, ..., \hat{p}$ lags. Analogously, h-step ahead forecasts of the factors are produced by:

$$\hat{\boldsymbol{F}}^k_{t+h|t} = \left(\hat{\boldsymbol{\beta}}^k_{\hat{p}}\right)^h \hat{\boldsymbol{F}}^k_t \tag{11}$$

When exogenous variables are considered for the prediction of the factors, $\hat{\boldsymbol{F}}^k_t$ in equation 11 contains the factors created by PCA and the exogenous variables. Also the coefficients for the exogenous variables are included in coefficients so that $\hat{\boldsymbol{\beta}}^k_{\hat{p}} = [c^k, \boldsymbol{\Phi}^k_j, \boldsymbol{\gamma}^k]$ for $j = 1, ..., \hat{p}$.

Since the IVS of equity options on stock k at day $t + h$ given the information up to day t depends only on $\hat{\boldsymbol{F}}^k_{t+h|t}$, forecasts of the IVS can be produced by:

$$\boldsymbol{I}\hat{\boldsymbol{V}}^k_{t+h|t} = \hat{\boldsymbol{\Theta}}^k_t \hat{\boldsymbol{F}}^k_{t+h|t} = \hat{\boldsymbol{\Theta}}^k_t \left(\hat{\boldsymbol{\beta}}^k_{\hat{p}}\right)^h \hat{\boldsymbol{F}}^k_{t|t} \tag{12}$$

The factor loadings $\boldsymbol{\Theta}^k_t = (\boldsymbol{\theta}^k_{1,t}, \boldsymbol{\theta}^k_{2,t}, ..., \boldsymbol{\theta}^k_{r,t})'$ are time-varying since for each rolling window the selected factors/exogenous variables by LARS may vary.

2.1.5 Algorithm

To summarise, the two-step approach to forecast $\hat{\boldsymbol{IV}}^k_{t+h|t}$ is as follows (repeat the steps for every rolling window):

1. Compute the factors by PCA and fit the IVS based on the factor decomposition:

$$\boldsymbol{IV}^k_t = \boldsymbol{\theta}^k_1 F^k_{1,t} + \cdots + \boldsymbol{\theta}^k_r F^k_{r,t} + \boldsymbol{\epsilon}_t = \boldsymbol{\Theta}^k \boldsymbol{F}^k_t + \boldsymbol{\epsilon}_t \tag{13}$$

   Model the time-variation of the IVS with the VAR model for the factors.

$$\hat{\boldsymbol{F}}^k_t = \boldsymbol{c}^k + \sum_{j=1}^{p} \boldsymbol{\Phi}^k_j \hat{\boldsymbol{F}}^k_{t-j} + \boldsymbol{v}_t \qquad \boldsymbol{v}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}^k_{\hat{\boldsymbol{F}}^k \hat{\boldsymbol{F}}^k}) \tag{14}$$

   To improve the forecasts of the factors extend the VAR model with exogenous variables, $\boldsymbol{X}_t = [\boldsymbol{r}^{SPX}_t, \boldsymbol{r}^{SPX}_{t-1}, \boldsymbol{I}_t, \boldsymbol{\pi}_t, \boldsymbol{Y}_t, \boldsymbol{F}^{SPX}_t]$.

$$\hat{\boldsymbol{F}}^k_t = \boldsymbol{c}^k + \sum_{j=1}^{p} \boldsymbol{\Phi}^k_j \hat{\boldsymbol{F}}^k_{t-j} + \boldsymbol{\gamma}^k \boldsymbol{X}_t + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \boldsymbol{\Omega}) \tag{15}$$

   Select the most relevant factors for the prediction of the IVS using the LARS algorithm. The factors created by PCA and the exogenous variables are used as input for the LARS algorithm. Select the number of r factors that gives the best prediction.

2. The h-step ahead forecasts of the factors selected by LARS are produced by:

$$\hat{\boldsymbol{F}}^k_{t+h|t} = \left(\hat{\boldsymbol{\beta}}^k_{\hat{p}}\right)^h \hat{\boldsymbol{F}}^k_t \tag{16}$$

   where $\hat{\boldsymbol{\beta}}^k = [\hat{c}^k, \hat{\boldsymbol{\Phi}}^k_j]$ are the coefficients of the projection of $\hat{\boldsymbol{F}}^k_t$ on $\hat{\boldsymbol{F}}^k_{t-h}$ and its $\hat{p}$ lags. When exogenous variables are considered for the prediction, $\hat{\boldsymbol{F}}^k_t$ in equation 16 contains the factors created by PCA and the exogenous variables. Also the coefficients for the exogenous variables are included in coefficients such that $\hat{\boldsymbol{\beta}}^k_{\hat{p}} = [c^k, \boldsymbol{\Phi}^k_j, \boldsymbol{\gamma}^k]$ for $j = 1, ..., \hat{p}$.

   The h-step ahead forecasts of the IVS of equity options can be produced by:

$$\hat{\boldsymbol{IV}}^k_{t+h|t} = \hat{\boldsymbol{\Theta}}^k_t \hat{\boldsymbol{F}}^k_{t+h|t} = \hat{\boldsymbol{\Theta}}^k_t \left(\hat{\boldsymbol{\beta}}^k_{\hat{p}}\right)^h \hat{\boldsymbol{F}}^k_{t|t} \tag{17}$$

   The factor loadings $\boldsymbol{\Theta}^k_t = (\boldsymbol{\theta}^k_{t,1}, \boldsymbol{\theta}^k_{t,2}, \cdots, \boldsymbol{\theta}^k_{t,r})'$ are time-varying since for each rolling window the selected factors/exogenous variables by LARS may vary.

2.2 One-step approach

The alternative to the two-step approach described in section 2.1 is to estimate the factors and the IVS simultaneously. van der Wel et al. [2016] describe the model for the implied volatility as state space model, where the factors are treated as unobserved latent factors. A state space model consist of two equations:

the observation equation that links the observed implied volatilities series for N different options on stock k to the states (factors), and the state equation that describes how the factors change over time. The representation of this model is indicated below, where equations 18 and 19 denote the observation and state equation respectively:

$$\boldsymbol{IV}_t^k = \boldsymbol{\Theta}^k \boldsymbol{F}_t^k + \boldsymbol{w}_t \qquad \boldsymbol{w}_t \sim N(0, \boldsymbol{R}) \tag{18}$$

$$\boldsymbol{F}_{t+1}^k = \boldsymbol{\Phi}^k \boldsymbol{F}_t^k + \boldsymbol{v}_{t+1} \qquad \boldsymbol{v}_{t+1} \sim N(0, \boldsymbol{Q}) \tag{19}$$

In these equations, $\boldsymbol{IV}_t^k$ is the observation vector with the log implied volatilities for each option on stock k, $\boldsymbol{F}_t^k$ are the unobserved $(r \times 1)$ factors, and $\boldsymbol{\Theta}^k$ is the $(N \times r)$ factor loading matrix. According to the approach of van der Wel et al. [2016], I model the state equation as a vector autoregressive model with $p = 1$ lag, where $\boldsymbol{\Phi}^k$ is a $(r \times r)$ loading matrix. It is possible to include more lags, however this is found to be computationally costly and $p = 1$ yields superior results. Finally, $\boldsymbol{R}$ and $\boldsymbol{Q}$ represent the $(N \times N)$ and $(r \times r)$ covariance matrices of the error processes in equation 18 and equation 19 respectively. The disturbances are mutually and serially uncorrelated. It is assumed that they follow a normal distribution, however this is not a requirement for the properties of the Kalman filter to hold. The system is stable if the eigenvalues of $\boldsymbol{F}_t^k$ lie in the unit circle, then $\boldsymbol{F}_t^k$ is mean-reverting to zero [Hamilton, 1994].

The unobserved factors $\boldsymbol{F}_t^k$ can be estimated by the Kalman filter. The algorithm was introduced by Kalman [1960] for engineering applications, however it is also used in other disciplines. The forecasting exercise using the Kalman filter proposed in this paper is mainly based on Hamilton [1994]. The Kalman filter is a recursive procedure of two different steps: prediction and updating, through the time index $t \in [0, T]$. It provides an optimal forecast of the factors $\boldsymbol{F}_t^k$ given the information at time t. There also exist a Kalman smoother, which smooths the estimates of the factors conditional on the complete sample size by backward recursion from $T - 1$ to 1. However, the Kalman filter requires the parameters in equations 18 and 19: $\boldsymbol{\Theta}^k$, $\boldsymbol{R}$, $\boldsymbol{\Phi}^k$ and $\boldsymbol{Q}$ to be known to construct estimates of the factors. The parameters can be estimated by Maximum Likelihood (ML) but with many states it may converge to a local minimum. The Expectation Maximisation (EM) algorithm of Shumway and Stoffer [1982] is more robust since the maximisation step is done analytically which is not possible to do when using ML. The EM-algorithm also ensures that the covariance matrices $\mathbf{R}$ and $\mathbf{Q}$ are positive definite and symmetric. The aim of the EM algorithm is to maximise the following joint density of observing the states $\boldsymbol{F}_t^k$ and the data $\boldsymbol{IV}_t^k$:

$$\begin{aligned} \mathrm{L}(\boldsymbol{IV}_t^k, \boldsymbol{F}_t^k | \boldsymbol{\theta}) = \frac{T}{2}\log|\boldsymbol{R}^{-1}| - \frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{IV}_t^k - \boldsymbol{\Theta}^{k'}\boldsymbol{F}_t^k)'\boldsymbol{R}^{-1}(\boldsymbol{IV}_t^k - \boldsymbol{\Theta}^{k'}\boldsymbol{F}_t^k) \\ + \frac{T}{2}\log|\boldsymbol{Q}^{-1}| - \frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{F}_t^k - \boldsymbol{\Phi}^k\boldsymbol{F}_{t-1}^k)'\boldsymbol{Q}^{-1}(\boldsymbol{F}_t^k - \boldsymbol{\Phi}^k\boldsymbol{F}_{t-1}^k) \end{aligned} \tag{20}$$

where the optimisation is for $\boldsymbol{\theta} = [\boldsymbol{\Theta}^k, \boldsymbol{R}, \boldsymbol{\Phi}^k, \boldsymbol{Q}]$. The EM algorithm steps are as follows: first it takes some initial values for the states, runs the Kalman filter with the values from 1 to T and smooths the estimates backward from $T-1$ to 1. It takes the expectation of the states conditional on the complete information set, which are the result of the Kalman smoother. Subsequently the algorithm maximises the resulting likelihood function given in equation 20 with respect to the unknown parameters $\boldsymbol{\theta}$. The parameter estimates are used again in the Kalman filter to predict the states. The EM algorithm iterates the expectation and maximisation steps until the parameters are converged. A more detailed description of the Kalman filter and the expectation maximisation algorithm is found in appendix A.1.3.

For identification, I impose the following restrictions[1] inspired by van der Wel et al. [2016]: $\boldsymbol{R}$ is a diagonal matrix, which means that there is no correlation between the observations equations. Also, the diagonal elements of $\boldsymbol{Q}$ are equal to 1, which imposes that the variances of the disturbances of the state equation are equal to 1. I initialise the algorithm using the first 200 observations of the data: $\boldsymbol{\Theta}^k$ is the coefficient matrix of the principal component analysis and $\boldsymbol{R}$ as the corresponding residuals. $\boldsymbol{\Phi}^k$ is initialised as the coefficient matrix of a VAR(1) model of the first r factors created by PCA, and $\boldsymbol{Q}$ as the covariance of the VAR(1) model.

Subsequently, forecasting in the state space framework is straightforward. The forecasts of the observation $\hat{\boldsymbol{IV}}^k_{t+h}$ conditional on all available observations $\boldsymbol{IV}^k_1, \ldots, \boldsymbol{IV}^k_t$, is denoted by:

$$\hat{\boldsymbol{IV}}^k_{t+h|t} = \boldsymbol{\Theta}^k \hat{\boldsymbol{F}}^k_{t+h|t} = \boldsymbol{\Theta}^k (\boldsymbol{\Phi}^k)^h \hat{\boldsymbol{F}}^k_{t|t} \tag{21}$$

Given the characteristics of the Kalman filter there is no need to use a rolling window for estimation. The Kalman filter is recursive, therefore the current estimated factors depend on the previous factors. This implies that the future values of the unobserved factors $\boldsymbol{F}^k_{t+1}$ depend on the current value $\boldsymbol{F}^k_t$ and thus indirectly on $(\boldsymbol{F}^k_{t-1}, \boldsymbol{F}^k_{t-2}, ..., \boldsymbol{F}^k_0)$. In this way, the Kalman filter efficiently uses the information from the implied volatility and the factors up to time t to predict $t+1$.

Similar to the two-step approach I extend the state-space model with exogenous variables. The same set of exogenous variables $\boldsymbol{X}_t = [\boldsymbol{r}^{SPX}_t, \boldsymbol{r}^{SPX}_{t-1}, \boldsymbol{I}_t, \boldsymbol{\pi}_t, \boldsymbol{Y}_t, \boldsymbol{F}^{SPX}_t]$ as in the two-step approach is considered. There is no algorithm to select state-space factors, since these are typically unobserved. Therefore the LARS algorithm cannot be used to select the exogenous variables along the state space factors. The exogenous variables selected in the two-step approach, are also used to extend the state-space model in the one-step approach. The state space model then becomes:

$$\boldsymbol{IV}^k_t = \boldsymbol{\gamma}^k \boldsymbol{X}^k_t + \boldsymbol{\Theta}^k \boldsymbol{F}^k_t + \boldsymbol{w}_t \qquad \boldsymbol{w}_t \sim N(0, \boldsymbol{R}) \tag{22}$$

$$\boldsymbol{F}^k_{t+1} = \boldsymbol{\Phi}^k \boldsymbol{F}^k_t + \boldsymbol{v}_{t+1} \qquad \boldsymbol{v}_{t+1} \sim N(0, \boldsymbol{Q}) \tag{23}$$

Therefore $\boldsymbol{F}^k_{t+1}$ indirectly depends on the exogenous variables $\boldsymbol{X}_t$ through the observation $\boldsymbol{IV}^k_t$. Since $\boldsymbol{X}_t$

[1] In unreported results I have tried different sets of restrictions, but these restrictions have proven to give the best results

is deterministic, the factors contain all the information that is relevant for predicting the IVS:

$$\begin{aligned}\hat{\boldsymbol{IV}}^k_{t+h|t} &= \mathrm{E}(\boldsymbol{\gamma}^k \boldsymbol{X}^k_{t+h} + \boldsymbol{\Theta}^k \boldsymbol{F}^k_{t+h} + \boldsymbol{w}_{t+h}|\boldsymbol{F}_t, \boldsymbol{F}_{t-1} \cdots, \boldsymbol{IV}^k_t, \boldsymbol{IV}^k_{1-1}, \cdots) \\ &= \boldsymbol{\gamma}^k \boldsymbol{X}^k_{t+h} + \boldsymbol{\Theta}^k \mathrm{E}(\boldsymbol{F}^k_{t+h}|\boldsymbol{F}_t, \boldsymbol{F}_{t-1} \cdots, \boldsymbol{IV}^k_t, \boldsymbol{IV}^k_{1-1}, \cdots) \\ &= \boldsymbol{\gamma}^k \boldsymbol{X}^k_{t+h} + (\boldsymbol{\Theta}^k)^h \boldsymbol{F}^k_t\end{aligned}$$

2.3 Benchmark model

For both testing and comparative purposes, I also estimate two benchmark models. Both benchmarks are partly equal to the the two-step approach in section 2.1. For the first benchmark model I will use the parametric model for the IVS of Goncalves and Guidolin [2006]. The parametric model is similar to the two-step approach but with a deterministic factor set. These factors are called the deterministic factors since all variables are exogenous and option characteristics.

$$\boldsymbol{IV}^k_t = \boldsymbol{\iota}_N \beta^k_{0,t} + \boldsymbol{m}^k \beta^k_{1,t} + \boldsymbol{m}^{k^2} \beta^k_{2,t} + \boldsymbol{\tau}^k \beta^k_{3,t} + (\boldsymbol{m}^k \boldsymbol{\tau}^k)\beta^k_{4,t} + \boldsymbol{\epsilon}_t = \boldsymbol{\Theta}^k_{DF} \boldsymbol{F}^k_{DF,t} + \boldsymbol{\epsilon}_t \tag{24}$$

where $\boldsymbol{\epsilon}_t$ is the random error term (assumed to be white noise). $\boldsymbol{m}$ and $\boldsymbol{\tau}$ refer to $(N \times 1)$ column vectors containing respectively the moneyness and time-to-maturity, for each option of stock k. $\boldsymbol{\Theta}^k_{DF} = [\boldsymbol{\iota}_N, \boldsymbol{m}^k, \boldsymbol{m}^{k^2}, \boldsymbol{\tau}^k, (\boldsymbol{m}^k \boldsymbol{\tau}^k)]$ is the $(N \times 5)$ matrix containing the explanatory variables which do not vary over time. The univariate time-series are tested for unit-root. Estimating equation 24 with OLS for each day results in a five-dimensional time series $\hat{\boldsymbol{F}}^k_{DF,t}$, which contains OLS coefficients with the following interpretation, $\hat{\boldsymbol{\beta}}^k_{0,t}$ is the intercept/level coefficient, $\hat{\boldsymbol{\beta}}^k_{1,t}$ is the moneyness slope of the IVS, $\hat{\boldsymbol{\beta}}^k_{2,t}$ describes the curvature of the IVS in the moneyness dimension, $\hat{\boldsymbol{\beta}}^k_{3,t}$ is the maturity slope, and $\hat{\boldsymbol{\beta}}^k_{4,t}$ captures possible interactions between the moneyness and time-to-maturity. To create predictions of the IVS using the factor decomposition in equation 24 similar steps are taken as described in subsection 2.1.5. The time-variation is modeled using a VAR model as in equation 14. The prediction of the IVS is done using equation 17. This benchmark model is also extended with the set of exogenous variables $\boldsymbol{X}_t = [\boldsymbol{r}^{SPX}_t, \boldsymbol{r}^{SPX}_{t-1}, \boldsymbol{I}_t, \boldsymbol{\pi}_t, \boldsymbol{Y}_t, \boldsymbol{F}^{SPX}_t]$.

As a second benchmark model I consider the ad-hoc Strawman model of Dumas et al. [1998]. The model fits a random walk to the IVS coefficients in the deterministic model in equation 24. This implies that the best forecast h-step ahead is the estimate of today [Bernales and Guidolin, 2014, Goncalves and Guidolin, 2006].

$$\hat{\boldsymbol{F}}^k_{DF,t+h|t} = \hat{\boldsymbol{F}}^k_{DF,t} + \boldsymbol{v}_t \qquad \boldsymbol{v}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}^k_{\hat{\boldsymbol{F}}^k_{DF}\hat{\boldsymbol{F}}^k_{DF}}) \tag{25}$$

This corresponds to a special case of equation 14 with $\boldsymbol{c}^k = 0$, $\hat{p} = 1$, $\boldsymbol{\Phi}_1 = \boldsymbol{I}_r$. The same steps as described in subsection 2.1.5 are followed to create forecasts of the equity IVS.

2.4 Evaluation measures

The performance for out-of-sample forecasting for each approach is evaluated based on three measures. I will report the mean of the following three statistics over time and across equity options for each approach $a \in$ [benchmark, one-step, two-step].

- The root mean squared forecast error (RMSE) which is the squared deviation between the actual and forecasted implied volatility of the options on stock k by approach a at day t:

$$\text{RMSE}_{a,t}^{k} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{IV}_{i,t+h}^{k} - \boldsymbol{I\hat{V}}_{i,t+h|t}^{k,a})^2} \tag{26}$$

- The mean absolute forecast error (MAE) i.e. the average absolute deviation between the actual and forecasted implied volatility of the options on stock k by model a at day t:

$$\text{MAE}_{a,t}^{k} = \frac{1}{N}\sum_{i=1}^{N}|\boldsymbol{IV}_{i,t+h} - \boldsymbol{I\hat{V}}_{i,t+h|t}^{k,a}| \tag{27}$$

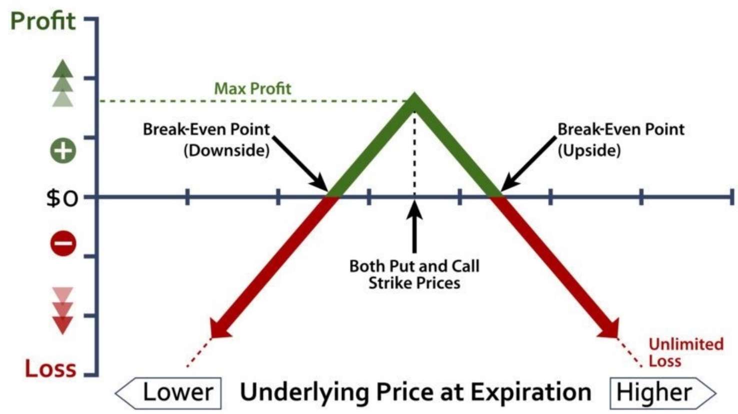
- The mean correct prediction of direction of change (MCPDC) which is defined as the percentage of predictions of the options on stock k for which model a correctly forecasted the direction of the change for the next day $t+1$:

$$\text{MCPDC}_{a,t}^{k} = \frac{1}{N}\sum_{i=1}^{N}1_{[\text{sign}(\boldsymbol{IV}_{i,t+h}^{k} - \boldsymbol{IV}_{i,t}^{k}) = \text{sign}(\boldsymbol{I\hat{V}}_{i,t+h|t}^{k,a} - \boldsymbol{IV}_{i,t}^{k})]} \tag{28}$$

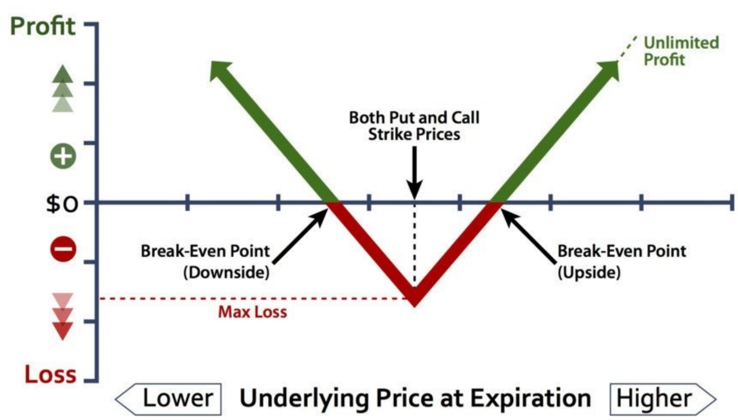
To see if whether the differences in forecasting performances between the approaches are signicant, I perform the equal predictive ability test proposed by Diebold and Mariano [2002]. The test has a null-hypothesis of equal forecast accuracy given a preferred loss function. Since there is no reason to assume that positive and negative errors have a different impact, I use equal performance as null hypotheses. In particular, I compare the forecasting performances of the approaches based on the above described performance measures to those of the random walk model. I illustrate this with an example, I test if the differences between the RMSE statistics of the two-step approach and the RMSE statistics of the random walk model are significant. For significant positive values the null hypothesis is rejected in favor of the considered approach. In case of negative values, the hypothesis is rejected in favor of the random walk model.

2.5 Trading strategy

The economic value of IVS forecasts is assessed by constructing a trading strategy based on the one-step ahead forecasts of the approaches. Following Bernales and Guidolin [2014] I generate trading portfolios based on an ATM straddle strategy. The advantage of a straddle strategy is that it is free of risks caused by changes in the prices of the underlying stocks. Delta-hedged strategies are also free of risk caused by changes in the underlying, but require investing in or borrowing underlying shares for which none of the models in this paper are constructed. The straddle strategy takes the full advantage of predictability patterns in implied volatilities, since it trades only equity options [Bernales and Guidolin, 2014].



(a) Short Straddle



(b) Long Straddle

Figure 3: The payoff of a straddle strategy. Figure 3a shows the payoff structure for a short straddle position and figure 3b shows the payoff structure for a long straddle position.

A straddle strategy involves trading a combination of a call and a put option with the same strike price and expiration date. Straddles are a good strategy to execute if a trader believes that the price of the stock will move significantly but he is not sure in which direction. The trader makes a profit if the stock price moves up or down enough to cover the premium paid. The premium consists of the price for the puts and calls plus the transaction costs. In figure 3 the payoff structure for a long and short straddle position are presented. A long straddle position earns profits when the price of the underlying stock is larger than the strike price of the long call option plus the premium paid or the price of underlying stock is smaller than the strike price of the long put minus the premium paid. The maximum possible loss is the premium paid. This maximum loss occurs when the price of the underlying stock equates the strike price of the long call/put. Therefore if high volatility is expected the straddle strategy is a suitable strategy. A long straddle, in which call and put options are purchased, is equivalent to a bet on a high(er) future volatility. A short straddle, in which call and put options are sold, is equivalent to a bet on a low(er) future volatility. Different performance for each maturity-moneyness groups is expected, therefore if all groups would be considered at once it would not be possible to draw a conclusion about the economic value of the forecasts based on the results. For this reason I focus only on ATM straddles. An option is said to be ATM if the current price of the underlying stock and the strike price of the option are equal. An ATM option has no intrinsic value, only time value. This means it would not make money if it were to expire today.

On every day I choose an ATM straddle for every stock in which I invest a fixed amount of \$1,000. The implied volatilities of those puts and calls are predicted using the different approaches. The straddle is purchased (sold) when the model forecasts that the implied volatility for equity options on stock k will increase (decrease) between t and $t+1$. At day $t+1$ the profits and losses of the positions are recorded and again \$1,000 for every stock is invested in the straddles. This trading strategy is applied for every day in the out-of-sample dataset.

I first explain the strategy if no transaction costs are considered. If the implied volatility of both the call and put option in the straddle on stock k is predicted to increase between day t and $t+1$ the options are purchased. I purchase the quantity $X_t^k = \$1000/(C_{m,t}^k + P_{m,t}^k)$ in units of straddles on stock k, for a total cost of \$1000. Let $C_{m,t}^k$ ($P_{m,t}^k$) denote the call (put) price of an option on stock k with maturity m at day t. The options only differs across maturities, since a straddle implies that the options have the same strike price and expiration date. Consequently, the profit made by the trading is as follows:

$$\text{Profit}_{t+1}^k(\text{up}) = X_t^k \big[\sum ((C_{m,t+1}^k + P_{m,t+1}^k) - (C_{m,t}^k + P_{m,t}^k)) \big] \tag{29}$$

However, when the forecasted implied volatility of the straddle on stock k of the portfolio is negative, I sell the quantity $X_t^k = \$1000/(C_{m,t}^k + P_{m,t}^k)$ in units of the straddle portfolio. This generates a cash inflow of \$1000. I also invest the \$1000 generated by selling the straddles together with the \$1000 initially investment at the riskless interest rate r_t over one day. In this case, the profit is:

$$\text{Profit}_{t+1}^k(\text{down}) = \$2000 * (\exp(\frac{r_t}{252}) - 1) \tag{30}$$

If there does not exist a straddle at day t on stock k, I assume that the trader invests \$1,000 in the riskless asset for one trading day. In this case $\text{Profit}_{t+1}^k(\text{equal}) = \$1000 * (\exp(\frac{r_t}{252}) - 1)$.

I include a benchmark strategy for comparison. This benchmark based on a daily investment of \$1000 in the risk-free interest rate rolled over time, which only yields the time value of money. The associated one-day net profit is given by $\text{Profit}_{t+1}^{k}(\text{risk-free}) = \$1000 * (\exp(\frac{r_t}{252}) - 1)$.

Note that the above described trading strategy is simplified by assumptions regarding market liquidity and transactions costs. Therefore profits generated by the strategy mostly cannot be achieved in the real world trading. To make the trading strategy more realistic, transaction costs are accounted for. The considered costs are for every transaction a half times the bid-ask spread. The transaction costs are subtracted from the profits on a daily basis.

The economic value of the predictability of the IVS is evaluated based on the generated average daily profit and the corresponding standard errors for the different approaches. A two-tailed t-test will be performed to investigate whether the profits are significantly different from zero. In addition, to gain a better understanding of whether the approaches are useful for traders, Sharpe ratios are provided. The Sharpe ratio is the average return earned in excess of the risk-free rate over the total risk of the portfolio.

$$SR_p = \frac{\bar{r}_p - r_f}{\sigma_p} \tag{31}$$

where $\bar{r}_p$ and σ_p denote mean profit and standard deviation of the straddle portfolio, and r_f denotes the average risk-free rate. This helps to understand whether the performance of a strategy is the result of smart trading or having high exposure to risk. The greater a portfolio's Sharpe ratio, the better its risk-adjusted performance.

3 Data

For this study, I will use datasets of daily implied volatilities for options on stocks obtained from OptionMetrics. The data include daily closing bid and ask quotes, option type (call/put), volume, strike prices, expiration date, option delta (Δ), dividends paid on each underlying asset and implied volatility. OptionMetrics uses the midpoint of the bid price and offer price as the theoretical price for the option in the Black-Scholes formula to calculate the implied volatility using a numerical search technique.

I consider a set of equity options on the ten stocks with the largest market capitalisation in the IT sector in the S&P 500. The equity options are specified in table 6 in appendix A.2. The data set covers the period from September 1, 2004, until December 31, 2017. Days on which the market is closed due to holidays or other reasons are excluded, resulting in 3,357 daily observations of the implied volatility. All considered companies have continuous listing during the sample period and the equity options are actively traded. Additionally, I use data on the daily implied volatilities for European options on the S&P500 index and daily time-series of the following exogenous variables: index returns (S&P 500), interest rates, expected inflation and corporate/government bond yield spread. The glossary and definitions of the exogenous state variables can be found in table 7 in appendix A.2.

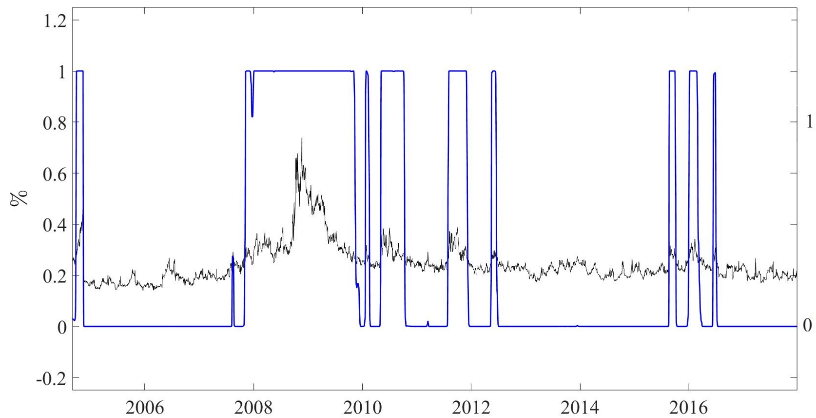


Figure 4: Markov-switching model for the implied volatility. The blue line corresponds to the estimated probabilities of a Markov-switching model mapped on the left axis. A probability of 0 corresponds to a low volatility regime for the implied volatility. A probability of 1 on the left axis corresponds to a high volatility regime for the implied volatility. Along the estimated probabilities the time-series of the implied volatilties averaged across the equity options is plotted in black corresponding to the right axis.

The datasets contain the 2008 financial crisis. The financial crisis was a period of extreme volatility which may influence results. A regime-switching model, allows the behavior of the variables to depend on the regime. A Markov switching model is such a regime-switching model which estimates the probability of being in regime 1 or regime 2. If I estimate a Markov switching model for a high and low volatility regime for the implied volatility, roughly two switches between the regimes are seen in figure 4. This is a reason to believe that the regime before the crisis is quite similar to after the crisis. Therefore I exclude the financial crisis from the sample period. The period from February 1, 2007, until September 1, 2009, is omitted. This results in a dataset of 2,706 daily observations of implied volatilities.

The datasets are filtered following the procedures of Dumas et al. [1998], Bernales and Guidolin [2014], van der Wel et al. [2016]. First, observations with implied volatility greater than 70% and those with missing values for either the implied volatility or Δ are dismissed. Second, all the options with maturity less than 10 days or longer than 360 days are dismissed, since these options usually contain little information regarding IVS. Third, I exclude options with a price lower than $0.30 for equity options and $\$\frac{3}{8}$ for S&P 500 index options, to avoid the effects of price discreteness on the IVS shape. Fourth, I consider only out-of-the-money (OTM) calls and puts as these are more actively traded than in-the-money options. The OTM call (put) can be matched to an ITM put (call) via the put-call parity where the Δ of the call option is always one plus the Δ of the put option and the corresponding put-call pair should have the same implied volatility. Therefore studying OTM options indirectly encompasses the study of ITM options.

I divide the data into different maturity-moneyness groups in a similar way as van der Wel et al. [2016]. More specifically I divide the data into three maturity groups classified as short-term if $\tau < 60$ days, medium term if $60 \leq \tau \leq 180$ days and long-term if $\tau > 180$ days. Together with six moneyness groups of at-the-money, out-of-the-money, and deep out-of-the-money options for both call and put options. I consider put options with $-0.125 < \Delta < 0$ as deep out-of-the-money (DOTM Put), with $-0.375 < \Delta < -0.125$ as out-of-the-money (OTM Put), and with $-0.5 < \Delta < -0.375$ as at-the-money (ATM Put). Similarly, call options with $0.375 < \Delta < 0.5$ are at-the-money (ATM Call), $0.125 < \Delta < 0.375$ are labeled as out-of-the-money (OTM Call) and with $0 < \Delta < 0.125$ are deep out-of-the-money (DOTM Call). The combination of the three maturities and six moneyness groups provides 18 different groups. Within each maturity-moneyness group, I select the option that has the smallest Euclidean distance to the midpoint of both dimensions as a representative for this group. If there is no option within a group on a specific day, these can be seen as missing values. I generate those missing values by selecting a neighbor defined as the option with the smallest Euclidean distance to both midpoints across all other groups at that day [van der Wel et al., 2016]. This may cause some extra bias in the predictions. The maturity-moneyness groups for long-term deep out-of-the-money puts and calls are empty most frequently. The groups for short-term and medium-term, out-of-the-money and at-the-money, calls and puts are almost never empty. This is logical as an option is considered deep out-of-the-money if its strike price is significantly above (for a call) or significantly below (for a put) the current price of the underlying asset. A call (put) option becomes in-the-money if the strike price is below (above) the stock price. Typically, the deeper out-of-the-money the option is, the less intrinsic value the options has, and therefore the less it is traded. Options trading activity tends to be higher when options are at-the-money or a bit out-of-the-money.

The filtered datasets will be split into two parts: the first 200 observations (first rolling window) of the datasets are used as an estimation period for the models and the other 2,506 observations are used to produce out-of-sample forecasts. I take the log of the implied volatility as the dependent variable in the models to ensure that the predictions are non-negative.

3.1 Data statistics

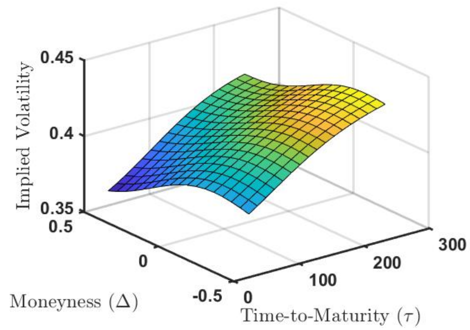
This subsection describes summary statistics and empirical characteristics of the implied volatility surface as a preliminary investigation of the IVS of equity options. In table 1 average data statistics across the equity options are presented. Reported are the mean and standard deviation of the implied volatilities, Δ and time-to-maturity. The statistics are ordered by the different maturity-moneyness groups. The implied volatility is quite flat but increases with time-to-maturity. An option with short time-to-maturity have a low implied volatility and options with long time-to-maturity tend to have a high implied volatility, since more time-value is priced into the option. The implied volatility is the highest for deep out-of-the-money puts and the lowest for deep out-of-the-money calls for each maturity. Δ is fairly constant across the maturities. The implied volatility is higher for puts than for calls for each group. Traders are typically more afraid of sharp drops in stock prices, than they are hopeful of strong rises. Table 8 in the appendix shows the average correlations and variances of the implied volatilities within the maturity-moneyness groups. Overall, high correlation between the maturity-moneyness groups is observed.

Table 1: Average data summary statistics across the equity options.

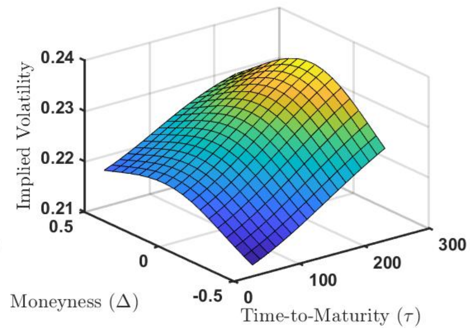
		$\tau < 60$		$60 < \tau < 180$		$\tau > 180$	
		Mean	Std	Mean	Std	Mean	Std
ATM call	IV	0.296	0.111	0.289	0.094	0.287	0.088
	Δ	0.225	0.216	0.198	0.215	0.196	0.220
	τ	55.689	54.140	91.395	54.432	145.612	91.719
OTM call	IV	0.288	0.107	0.282	0.089	0.285	0.086
	Δ	0.259	0.1425	0.264	0.104	0.293	0.117
	τ	45.882	30.018	126.562	27.369	222.167	71.623
DOTM call	IV	0.292	0.110	0.292	0.094	0.294	0.088
	Δ	0.398	0.134	0.428	0.084	0.427	0.116
	τ	48.420	38.897	125.915	31.788	214.603	75.469
DOTM put	IV	0.325	0.122	0.360	0.121	0.372	0.113
	Δ	0.141	0.240	0.022	0.210	0.029	0.212
	τ	54.070	50.422	113.824	47.103	201.018	87.483
OTM put	IV	0.282	0.120	0.328	0.108	0.334	0.101
	Δ	-0.207	0.210	-0.221	0.135	-0.130	0.210
	τ	43.935	27.001	126.920	26.295	222.698	71.260
ATM put	IV	0.299	0.113	0.302	0.099	0.304	0.092
	Δ	-0.263	0.309	-0.362	0.218	-0.285	0.300
	τ	48.839	39.572	125.309	32.268	214.535	76.000

Notes: the table shows data statistics ordered by moneyness and time-to-maturity. The mean and standard deviation (Std) are reported over time for the implied volatility (IV), Δ, and time-to-maturity τ (in days). The numbers represent averages over time for the selected equity options in each maturity-moneyness group.

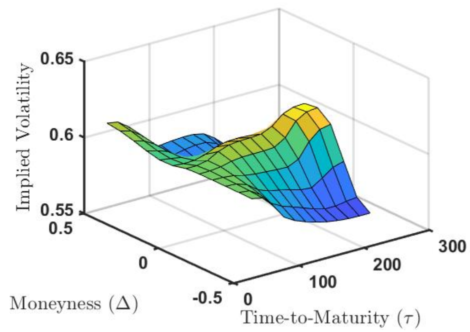
Figure 5 visualises the implied volatility surface of the options on the underlying stocks: Apple, NVIDIA,



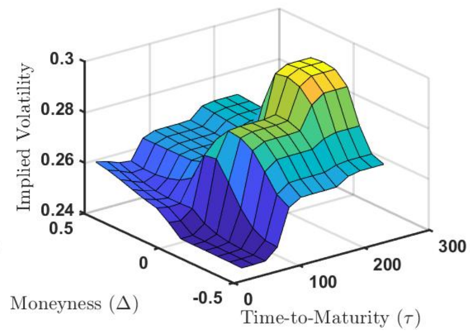
(a) Apple, August 20, 2008



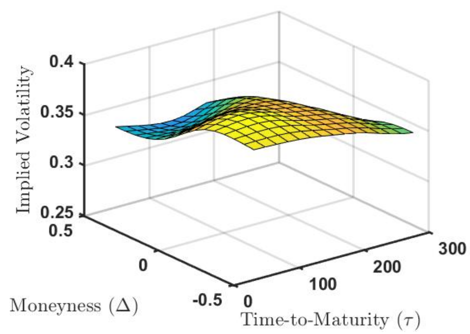
(b) Apple, September 1, 2014



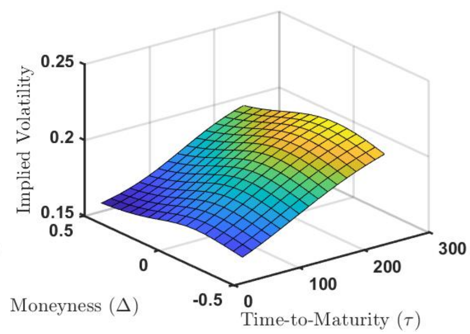
(c) NVIDIA, August 20, 2008



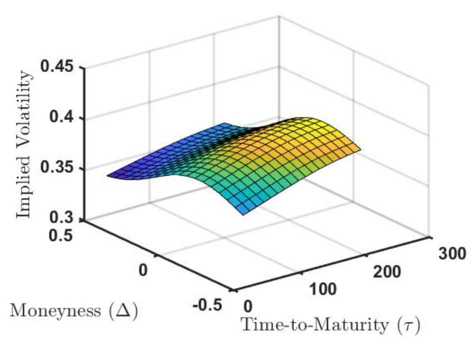
(d) NVIDIA, September 1, 2014



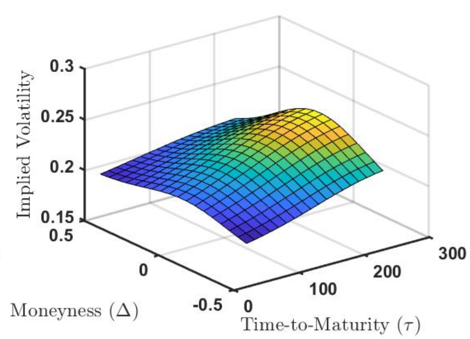
(e) Microsoft, August 20, 2008



(f) Microsoft, September 1, 2014



(g) Google, August 20, 2008



(h) Google, September 1, 2014

Figure 5: Equity IVS surfaces. Each subfigure corresponds to the IVS at a particular date for options on one of the underlying stocks. The color-mapped surface illustrates the fitted IVS. The surface is recovered from the 18 maturity-moneyness groups.

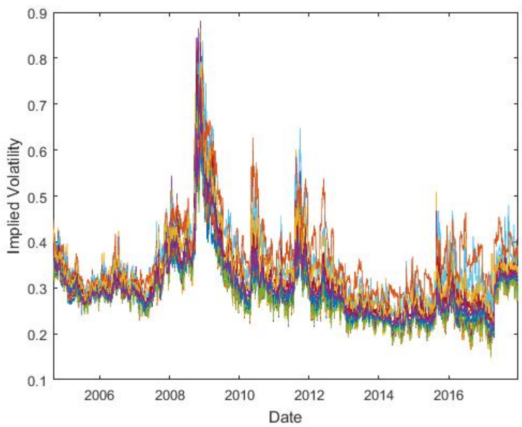
Microsoft and Google on two specific days. All the subfigures on the left show the IVS on August 20, 2008, during the financial crisis. The subfigures on the right show the IVS on September 1, 2014. The surfaces are recovered from the 18 maturity-moneyness groups. Apart from the level of the implied volatility, the shape of the IVS during the financial crisis is quite similar to the IVS after the financial crisis. Differences in the shape of the fitted surfaces can also be the results of different options that are selected for the groups on a specific day. Overall, the same features can be seen in the surfaces as in table 1. The implied volatility increases if the time-to-maturity increases and the implied volatility is higher for puts than for calls. However there are some exceptions, for example in figure 5e the implied volatilities slightly decrease with time-to-maturity. A common feature for implied volatility is the volatility smile. A volatility smile occurs in a situation in which at-the-money options have lower implied volatility than out-of-the-money. The volatility is higher in the extremes, deep out of the money and deep in the money, since the strike price is further away from the stock price. Therefore the shape of the IVS looks like a smile. This feature is also seen in figure 5. Around $\Delta = 0$, which corresponds to DOTM puts and DOTM calls, the implied volatility is the highest. The smile is not directly visible due to the ordering of the options. If the curve is more balanced to one side it is referred to as a volatility smirk. The sub-figures of figure 5 do not appear to show a relationship between the individual equities. There is no evidence that the IVS of two equities has involved in a similar way over time or that they have a comparable shape at the specific days.

Time series of the implied volatility over the sample period are shown in figure 6. Figure 6a gives the average time-series across the equity options for each maturity-moneyness group separately. The fluctuations are caused by the construction of the maturity-moneyness group. If there is no option in the group the option that is closest to the midpoints across all options is selected. Figure 6b displays the mean across the equity options and across all maturity-moneyness groups. It is clearly visible that the overall implied volatility is especially high during the financial crisis and has fewer peaks over the complete sample period. Figure 6c shows the average time-series across the maturity-moneyness groups for each individual equity option. There is variation in the level of the implied volatilities of the equity options. The evolution of the time-series is quite similar for each equity option. However, in the beginning and at the end of the sample period the time-series divert to some extent.

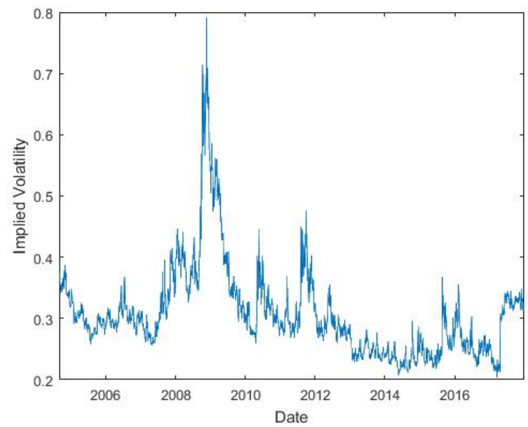
4 Empirical results

This section elaborates on the empirical study based on the data presented in section 3. I evaluate the predictions based on the two approaches outlined in section 2. The performance of the approaches will be compared to the benchmark models. I also assess the economic value of the out-of-sample forecasts by simulating the trading strategy described in section 2.5.

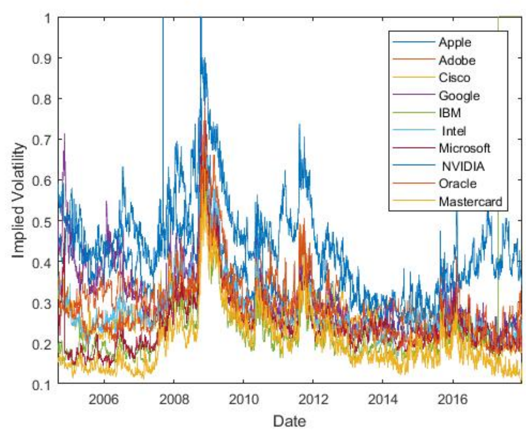
To ensure the right modeling approach, unit-root tests for the univariate times series of the factors and



(a) Average time-series across the equity options for each individual maturity-moneyness group



(b) Average time-series across equity options and maturity-moneyness groups



(c) Average time-series across the maturity-moneyness groups for each individual equity option

Figure 6: The time series of implied volatility over the sample period. Figure 6a gives the average time-series across the equity options for each individual maturity-moneyness group. Figure 6b the average time-series across equity options and maturity-moneyness groups. Figure 6c the average time-series across the maturity-moneyness groups for each individual equity option.

exogenous variables are performed. I test unit-root stationarity with the augmented Dickey-Fuller (ADF) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. The tests are performed on the univariate time-series for the complete out-of-sample length. In table 10 and table 11 in appendix A.3 respectively the ADF test statistics and the KPSS test statistics along with the corresponding p-values are reported for the factors created by PCA and the deterministic factors. The uniform conclusion from the ADF statistics is that the factors do not contain unit roots. The KPSS statistics are less uniform, however mostly the null-hypothesis of the absence of unit roots is not rejected. Table 9 shows the test statistics for the exogenous variables. Again no strong evidence for the presence of unit root is found. This implies that there is no need for modeling the log implied volatilities in first differences. Therefore the log implied volatility is modeled in levels.

4.1 Statistical results

In table 2 the average out-of-sample forecasting performance of the approaches for the implied volatility over time and across the equity options is presented. The out-of-sample dataset is from July 18, 2005, until December 29, 2017, excluding the financial crisis. The table reports the average root mean squared error (RMSE), the average mean absolute error (MAE) and the average mean correct prediction of the direction of change in implied volatilities (MCPDC) across the equity options. The corresponding Diebold-Marino (DM) statistics of equal predictability compared to the random walk are reported between brackets. The performance is evaluated for three forecast horizons: 1 day, 3 days and 5 days ahead. Panel A presents the results when no exogenous variables are considered. Panel B shows the performance when exogenous variables are considered. For the two-step approach, the most relevant predictors are selected by LARS from the aggregate set of factors created by PCA and the exogenous variables. For the benchmark model and the one-step approach, the exogenous variables selected by LARS algorithm in the two-step approach are used. The bold values indicate the best performing model for each forecasts horizon according to the performance measures.

When no exogenous variables are considered, the benchmark model is not able to outperform the random walk model in terms of RMSE and MAE for the short forecast horizons. However, in terms of MCPDC statistics the benchmark model performs significantly better compared to the random walk. When the horizon is increasing the difference between the two becomes smaller. Both the two-step and one-step approach significantly outperform the benchmark model and the random walk model in terms of all three measures. This confirms that more sophisticated models are better able to capture the dynamics of the IVS of equity options. The performance of the one-step approach is significantly higher than the performance of the two-step approach. This gives an indication that the IVS prediction accuracy is higher when the factors are estimated simultaneously with the IVS. In general the forecast accuracy diminishes as the forecast horizon increases. Both the one-step and two-step approach are able to correctly predict more than 50 percent of the directions of change in implied volatility (MCPDC statistics is higher than 0.5). If there were no predictability the methods are expected to be correct in at least 50% of the cases (MCPDC statistics is 0.5). This indicates that it is possible to create trading profits using the forecasts of the approaches.

The consideration of exogenous variables improves the forecast accuracy, see panel B of table 2. However, the forecast accuracy does not increase notably when considering exogenous variables. Although the improvement in forecast accuracy is minor, it is encouraging due to the limited number of the exogenous variables

Table 2: Average forecast performance measures across equity options.

	RMSE			MAE			MCPDC		
	h = 1	h = 3	h = 5	h = 1	h = 3	h = 5	h = 1	h = 3	h = 5
Random walk	0.0369 (NA)	0.0390 (NA)	0.0406 (NA)	0.0270 (NA)	0.0289 (NA)	0.0305 (NA)	0.5123 (NA)	0.5121 (NA)	0.5149 (NA)
Panel A: no exogenous variables									
Benchmark	0.0424 (-2.86)	0.0420 (-1.69)	0.0420 (-0.90)	0.0307 (-3.25)	0.0304 (-1.97)	0.0306 (-1.48)	0.5190 (0.69)	0.5199 (1.04)	0.5195 (0.88)
Two-step	0.0197 (6.02)	0.0678 (-0.99)	0.0232 (6.79)	0.0156 (8.65)	0.0552 (-0.99)	0.0189 (10.09)	0.5366 (-3.840)	0.5383 (-3.47)	0.5372 (-3.121)
One-step	**0.0138** (6.46)	**0.0153** (6.93)	**0.0172** (7.18)	**0.0113** (10.13)	**0.0122** (10.79)	**0.0136** (10.97)	**0.6643** (-33.78)	**0.6508** (-28.96)	**0.6346** (-25.87)
Panel B: including exogenous variables									
Benchmark	0.0390 (-2.26)	0.0370 (-1.77)	0.0379 (-1.68)	0.0288 (-2.56)	0.0278 (-2.16)	0.0289 (-2.16)	0.5282 (0.54)	0.5294 (0.72)	0.5308 (0.47)
Two-step	0.0175 (5.80)	0.0200 (5.88)	0.0225 (5.88)	0.0132 (8.73)	0.0155 (8.71)	0.0176 (8.39)	0.5502 (-3.99)	0.5469 (-3.25)	0.5446 (-2.83)
One-step	**0.0077** (6.67)	**0.0116** (7.11)	**0.0147** (7.40)	**0.0057** (10.46)	**0.0083** (10.82)	**0.0104** (10.94)	**0.7646** (-21.75)	**0.7154** (-20.79)	**0.6829** (-18.98)

Notes: The table contains the average root mean squared error (RMSE), average mean absolute error (MAE) and the average mean correct prediction of the direction of change in implied volatilities (MCPDC) across the equity options. The corresponding Diebold-Marino (DM) statistics of equal predictability compared to the random walk are reported between brackets. The performance is evaluated for three forecast horizons: 1 day, 3 days and 5 days ahead. The different factors sets are indicated as follows: Benchmark indicates the deterministic factors set (benchmark model), Two-step represents the two-step approach where the factors are created by PCA and One-step indicates the one-step approach where a state space model for the implied volatilities is estimated using the Kalman filter. Panel A shows the results when no exogenous variables are used. Panel B contains the performance when exogenous variables are considered. For the two-step, the most relevant predictors are selected by the LARS algorithm from the aggregate set of factors created by PCA and the exogenous variables. For the benchmark model and the one-step approach, the exogenous variables selected by LARS algorithm in the two-step approach are used. The bold values indicate the best performing model for each forecasts horizon according to the performance measures.

considered. This suggests that the consideration of other exogenous variables not available for this study could further contribute to the prediction of the IVS of equity options. In table 3 the percentage of rolling windows in which the exogenous variables and the factors created by PCA are selected by the LARS algorithm. The aggregate of the exogenous variables and the factors created by PCA is used as the input for the LARS algorithm. For each rolling window a maximum of five variables is selected. It is observed that the LARS algorithm selects five variables for every rolling window. Therefore the percentages for each equity option sum up to 500%. The LARS algorithm does not necessarily selects the same factors as a common PCA. The first $r = 5$ factors would be selected by PCA based on the large eigenvalues. As these are not chosen by LARS for every rolling window, this is a signal that the factors that explain most of the variation in the data set are always the best predictors. This indicates that the selection of more relevant factors improves the forecast performance of statistical factor models such as PCA. Often, exogenous variables are selected in addition to the factors. The interest rate and the expected inflation are frequently selected as a variable. Across the equities, these variables are selected in respectively 39% and 24% of all rolling windows. The (lagged) return of the S&P 500 is found to be a less useful variable than literature suggested. The selection of the exogenous variables and factors is different among the equity options and for each rolling window. Therefore, it is not possible to make one general factor set that can be used for each equity option over the complete sample. Hence, it is concluded that using the LARS algorithm to select the factors created by PCA and the exogenous variable improves the forecast accuracy.

So far, the forecasting performances have been analysed by considering average statistics across the equity options over the full out-of-sample period. To gain more insights on particular features of individual equity options of the IT sector and in what way the approaches perform over the out-of-sample time period, the statistical measures for each equity option are plotted. Figure 7 plots the time series of RMSE statistics. The upper panel plots the RMSE time-series for the benchmark model. The middle panel shows the time-series for the RMSE statistics when forecasting using the two-step approach. The lower panel contains the RMSE statistics for the one-step approach. Similar plots for the MAE statistics are found respectively in figure 9 in appendix A.3. The time-series have a visible amount of peaks for each option and each approach. Most peaks correspond to company-specific events. The benchmark approach (top panel) is most sensitive to the high volatility periods; before and during the crisis high values are observed. From the figures it cannot be conclude that one of the approaches is the best for options of any particular stock. Given that the performance measures of the approaches vary over time for each individual equity option, it might be of interest to construct forecast combinations in order to improve overall forecasting performances. However, these forecast combinations are beyond the scope of this thesis.

Table 4 reports the average RMSE statistics for each of the individual equity options using the approaches. The reported values are the RMSE statistics of the one-step ahead forecasts of both excluding and including exogenous variables averaged over three subperiods: before, during and after the financial crisis. Bold

Table 3: Percentage of rolling windows in which the exogenous variables and the factors created by PCA are selected by the LARS algorithm.

	Apple	Adobe	Cisco	Google	IBM	Intel	Mastercard	Microsoft	NVIDIA	Oracle
Return S&P 500	-	-	-	-	-	-	-	-	-	1.40
Lagged Return S&P 500	-	-	-	0.68	-	-	-	-	-	2.95
Expected inflation	9.94	45.25	24.66	19.11	30.85	26.14	23.25	0.36	34.12	28.93
Interest rate	17.80	64.88	35.55	32.68	43.02	44.21	61.88	3.79	51.48	40.66
Yield	-	-	-	0.84	-	0.04	14.01	-	-	2.75
Factors of index options	-	-	0.04	1.32	0.40	0.12	9.98	-	-	0.92
Factor 1	34.00	67.64	49.80	38.19	54.95	57.02	57.64	10.34	54.67	46.69
Factor 2	49.48	54.35	45.53	51.84	53.43	54.59	54.35	23.90	52.19	46.29
Factor 3	62.53	42.70	45.73	59.02	49.32	49.44	40.70	41.42	49.44	47.01
Factor 4	57.38	34.48	44.17	54.67	42.98	44.25	35.48	55.31	45.65	49.40
Factor 5	51.04	60.73	56.94	54.95	58.54	58.86	52.73	61.85	59.06	56.15
Factor 6	44.09	67.68	62.01	58.54	63.81	61.61	66.68	56.34	60.30	59.58
Factor 7	38.55	27.33	47.25	42.74	35.91	36.83	22.33	41.30	35.75	42.18
Factor 8	40.46	17.08	33.40	31.60	20.75	29.41	13.08	38.35	23.34	28.85
Factor 9	38.11	11.09	21.19	27.02	13.93	19.31	8.09	39.86	15.04	20.67
Factor 10	26.06	4.91	12.13	16.80	10.97	10.77	9.91	43.18	10.93	12.01
Factor 11	13.69	1.56	9.42	6.46	9.30	6.11	3.56	39.11	5.31	7.70
Factor 12	8.86	0.32	7.30	2.71	5.87	1.04	10.32	28.41	1.56	4.03
Factor 13	5.83	-	4.55	0.60	4.39	0.24	8.91	13.09	0.76	1.76
Factor 14	2.08	-	0.32	0.16	1.28	-	4.01	2.59	0.12	0.08
Factor 15	0.12	-	-	0.08	0.32	-	3.09	0.80	0.28	-
Factor 16	-	-	-	-	-	-	-	-	-	-
Factor 17	-	-	-	-	-	-	-	-	-	-
Factor 18	-	-	-	-	-	-	-	-	-	-

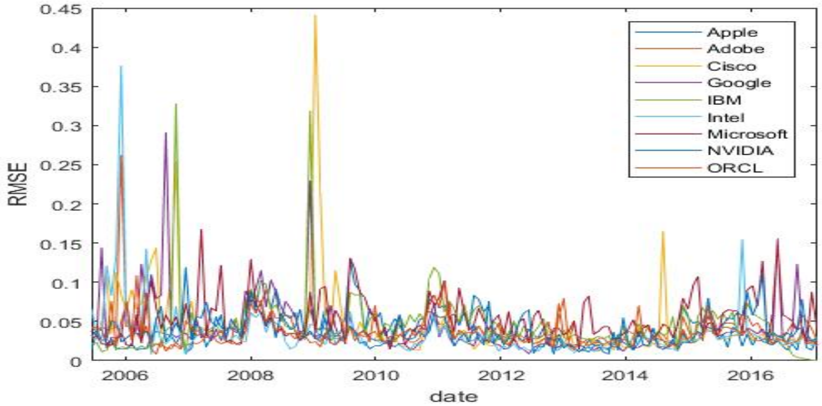
Notes: Percentage of rolling windows in which the exogenous variables and the factors created by PCA are selected by the LARS algorithm. The aggregate of the exogenous variables and the factors created by PCA is used as the input for the LARS algorithm. For each rolling window a maximum of five variables is selected. Factor 1 corresponds to the factor created by PCA with the largest eigenvalue. Likewise, Factor 18 corresponds to the factor created by PCA with the smallest eigenvalue. - indicates that the exogenous variable/factor is not selected by the LARS algorithm.

values represent the lowest RMSE statistic for the equity option in the specific sub-period. Similar tables containing the MAE and MCPDC statistics are found respectively in table 12 and table 13 in appendix A.3. The results of the approaches for three forecast horizons for each equity options when (no) exogenous variables are considered are found in table (15) 14 in appendix A.3. For each equity option it is observed that the RMSE statistics for the benchmark model differ in each subperiod. The performance of the two-step and the one-step approach is not highly influenced by the financial crisis. These statistics are steady over

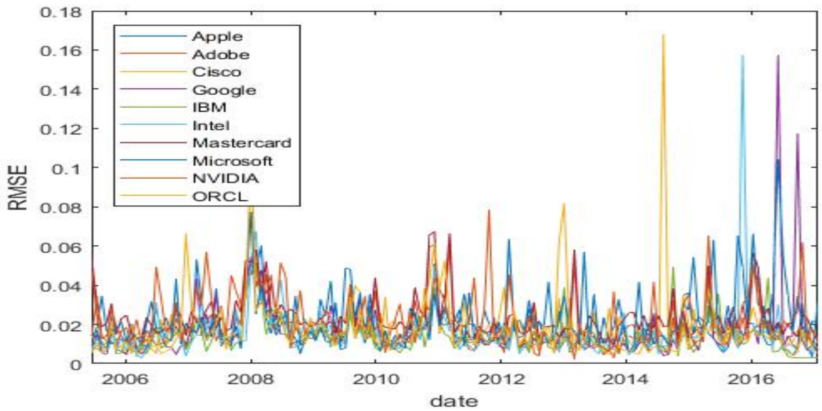
Table 4: Average RMSE statistics

		Before financial crisis *July 18, 2005 until January 31, 2007*	During financial crisis *February 1, 2007 until September 1, 2009*	After financial crisis *September 2, 2009 until December 31, 2017*
Apple				
	Benchmark	0.0340	0.0567	0.0376
	Two-step	0.0174	0.0263	0.0174
	One-step	**0.0106**	**0.0144**	**0.0098**
Adobe				
	Benchmark	0.0360	0.0458	0.0359
	Two-step	0.0142	0.0246	0.0160
	One-step	**0.0087**	**0.0168**	**0.0098**
Cisco				
	Benchmark	0.0952	0.0408	0.0264
	Two-step	0.0162	0.0262	0.0167
	One-step	**0.0089**	**0.0145**	**0.0103**
Google				
	Benchmark	0.0619	0.0649	0.0263
	Two-step	0.0136	0.0279	0.0143
	One-step	**0.0080**	**0.0162**	**0.0093**
IBM				
	Benchmark	0.0565	0.0561	0.0429
	Two-step	0.0114	0.0213	0.0149
	One-step	**0.0070**	**0.0120**	**0.0080**
Intel				
	Benchmark	0.0645	0.1739	0.0250
	Two-step	0.0139	0.0266	0.0145
	One-step	**0.0077**	**0.0147**	**0.0093**
Mastercard				
	Benchmark	0.0451	0.0503	0.0451
	Two-step	0.0205	0.0311	0.0223
	One-step	**0.0097**	**0.0239**	**0.0108**
Microsoft				
	Benchmark	0.0531	0.0915	0.0518
	Two-step	0.0208	0.0265	0.0231
	One-step	**0.0140**	**0.0151**	**0.0139**
NVIDIA				
	Benchmark	0.0351	0.0592	0.0351
	Two-step	0.0213	0.0332	0.0194
	One-step	**0.0133**	**0.0177**	**0.0115**
Oracle				
	Benchmark	0.0381	0.0408	0.0331
	Two-step	0.0149	0.0262	0.0165
	One-step	**0.0094**	**0.0145**	**0.0111**

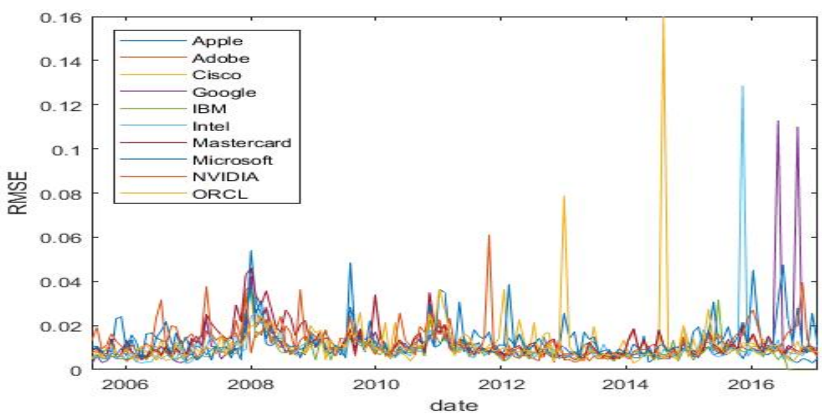
Notes: The table contains the average RMSE statistics of the one-step ahead prediction for the individual equity options. For each equity option the RMSE statistic for the three approaches is reported. The statistics are averaged over different subperiods: before, during and after the financial crisis. Bold values represents the lowest RMSE statistic for the equity option in the specific sub-period.



(a) Benchmark



(b) Two-step



(c) One-step

Figure 7: Time-series of the RMSE statistics for the individual equity options. The upper panel plots the time-series for the benchmark model. The middle panel plots the time-series for the RMSE statistics of the forecasts of the two-step approach. The lower panel shows the RMSE statistics for the one-step approach.

the subperiods. Overall, the one-step approach is the best performing model for each equity option in each subperiod according to the three performance measures.

4.2 The economic value of the predictions of equity IVS

The results in the previous section show the existence of predictability in the dynamics of the IVS of equity options. The predictability increases when exogenous variables are considered in the approaches. To assess the economic value of those predictions for traders a simple trading strategy as described in section 2.5 is constructed. The trading strategy tries to exploit the one-day-ahead predictions from the approaches to generate profit. The MCPDC statistics are highly correlated with the profits of the trading strategy if the simple trading strategy is followed: buy (sell) if the method predicts the IV will increase (decrease) between t and $t+1$. From section 4.1 it is concluded that the one-step approach including exogenous variables has the highest MCPDC statistics for one-step ahead prediction. This method is therefore expected to generate the highest profit.

Table 5 reports average daily summary statistics for the ATM straddle trading strategy applying the forecasts of the approaches. These are presented alongside the average summary statistics for effortless investment in the risk-free rate over the out-of-sample time period. Without taking transaction costs into account, the forecasts of all the approaches translate into significant positive average profits. All trading strategies, even the benchmark model, outperform the effortless investment in the risk-free rate. As expected the one-step approach yield the largest profit. However, it also has the highest standard errors, which means this is also the riskiest strategy. An average daily Sharpe ratio of 4.161 for the one-step approach indicates great performances. The Sharpe ratio of the risk-free investment is zero, since the expected return is the risk-free rate. Both the two-step approach and one-step approach have higher Sharpe ratio than the benchmark model and the effortless investment in the risk-free rate. It is concluded that, disregarding transaction costs, the approaches have sufficient economic value in forecasting the IVS of equity options.

However, when transaction costs are accounted for the profit of all approaches becomes negative as shown in panel B. The cost of each transaction is taken as half the bid-ask spread. The profit of the effortless investment in the risk-free rate remains positive since the transaction costs are negligible. The negative Sharpe ratios should be interpreted as that an investment in the risk-free rate would be more profitable than this strategy. However a negative Sharpe ratio can be misleading. When the mean profit is negative, the Sharpe ratio improves if the volatility is higher. This should be kept in mind when analysing negative Sharpe ratios. Relative to each other, the approaches show similar performances as when transaction costs were not accounted for. Again, the trading strategy based on the forecasts of the one-step approach is the least badly performing.

When exogenous variables are included in the approaches only minimal differences are seen in table 5 in

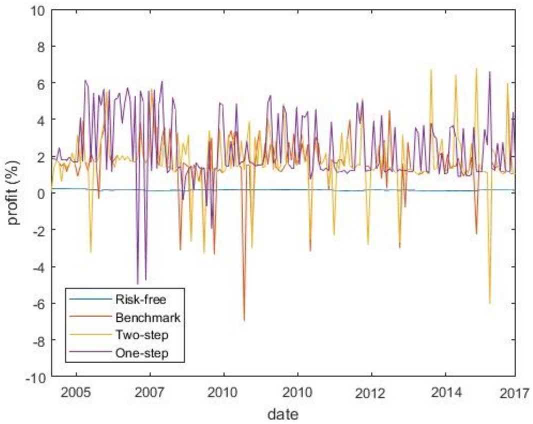
Table 5: Trading results of the ATM straddle trading strategy (before and after transaction costs)

	No exogenous variables				*Including obervable variables*			
	Mean Profit (%)	Std profit (%)	t-test	Sharpe ratio (%)	Mean Profit (%)	Std profit (%)	t-test	Sharpe ratio (%)
Before transaction cost								
Risk-free	0.300	0.241	62.263	0.000	0.300	0.241	62.263	0.000
Benchmark	3.123	7.200	21.718	2.733	3.601	7.474	24.121	3.035
Two-step	3.092	6.792	22.790	2.868	3.591	7.646	23.510	2.958
One-step	5.243	7.938	33.065	4.161	5.091	7.578	33.632	4.232
After transaction cost								
Risk-free	0.300	0.241	62.263	0.000	0.300	0.241	62.263	0.000
Benchmark	-3.538	7.542	-23.485	-2.955	-3.088	7.857	-19.678	-2.476
Two-step	-3.597	7.169	-25.114	-3.160	-3.191	8.016	-19.929	-2.508
One-step	-1.716	8.492	-10.115	-1.282	-1.826	8.136	-11.237	-1.424

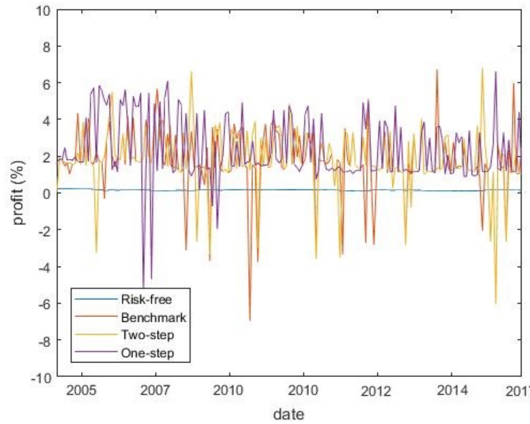
Notes: The table reports average daily summary statistics for the ATM straddle trading strategy applying the forecasts of the approaches. In particular, the average profit, the standard deviation of the profit, t-test statistic and the Sharpe ratio are provided as a daily percentage averaged over the full out-of-sample period. The ATM straddle trading strategy is applied for the forecasts of the approaches, indicated as: Benchmark, Two-step and One-step. In addition the results for an effortless investment in the risk-free rate over the entire time period are reported, indicated as: Risk-free. The statistics are presented in panel A when no transaction costs are accounted for, and in panel B transaction costs are included. The considered transaction costs taken for every transaction is half the bid-ask spread. The Sharpe ratios as defined in equation 31 are reported in percentages. A daily percentage of 4.161% translates into a $\frac{4.161}{100} * (252)^{(1/2)} = 0.6605$ annualised Sharpe ratio.

terms of mean profit and Sharpe ratios. This holds for both scenarios, before and after transactions costs. Remarkable is that the profits generated by the forecasts of the one-step approach diminish when exogenous variables are considered. Also the standard deviation of the one-step approach decreases, so that the Sharpe ratio improves minimally. For the trading strategy based on the forecasts of the two-step approach, the opposite is visible. The generated profits increase but the risk involved with the strategy also increases. From an economic perspective, including exogenous variables to the approaches is not very beneficial. However, the consideration of other exogenous variables not available for this study may contribute more positively to the economic value of the forecasts.

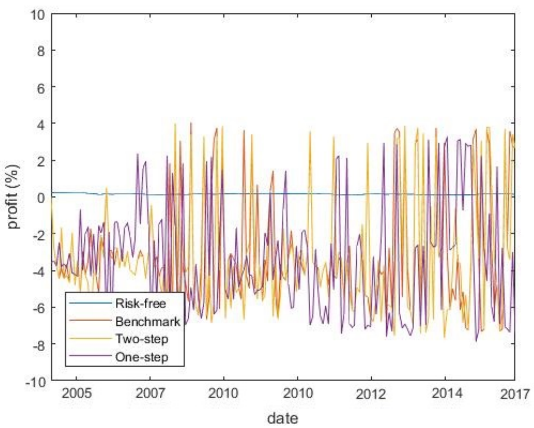
The dynamic performances of all trading strategies before and after including transaction costs are provided in figure 8. The approaches show similar patterns over time. The effortless investment in the risk-free



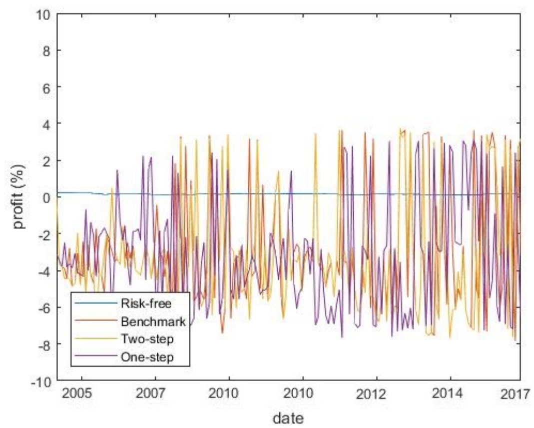
(a) No transaction costs



(b) Including exogenous variables, no transaction costs



(c) Including transaction costs



(d) Including exogenous variables and transaction costs

Figure 8: Time-series over the out-of-sample period of the profits created by the ATM straddle trading strategy applying the forecasts of the approaches. The upper panel, figure 8a and figure 8b, shows the performance when disregarding transactions. The lower panel, figure 8c and figure 8d, plots the profits when there transaction costs are accounted for.

rate has a constant performance. High negative profits over the full out-of-sample period are observed when transaction costs are considered. Similar to the findings of Bernales and Guidolin [2014], the results show that generating positive profits when accounting for transaction costs is difficult. Although the forecasts of the IVS of equity options for the different approaches have proven to be valuable from a statistical perspective, they are not able to generate profits using this trading strategy when transaction costs are considered.

4.3 Robustness

In this subsection, a number of robustness checks for the approaches are elaborated on. In section 3 it is seen that during the financial crisis the implied volatilities were very high. This period is excluded from the dataset. To determine how the consideration of this period would influence the findings, the models are rerun with the dataset including the financial crisis (unreported results). The performance of the models relative to one another does not change notably. The one-step approach is still the best performing method. The approaches were also applied using the data without a log-transformation. Overall the accuracy of the forecasts increases. However, the log-transformation is necessary for subsequent predictions since it prevents non-positive implied volatility forecasts.

5 Conclusion

The implied volatility surface is a mapping of implied volatilities of options as a function of the moneyness and time-to-maturity. Understanding the dynamics of these surface can contribute to trading and hedging strategies, as it contains information about the expected market volatility. This paper investigates whether the proposed two-step and one-step approach can provide accurate forecasts for the implied volatility surfaces of equity options. The two-step approach is based on factors created by a principal component analysis which are selected using the LARS algorithm. In the one-step approach, the factors are estimated simultaneously with the IVS. Both approaches are also applied with additional exogenous variables. The exogenous variables are selected based on the LARS algorithm for each rolling window. Only those exogenous variables that contribute to the performance of the models are used. The statistical and economic relevance of the approaches are compared.

There are three main findings in this paper. First, the approaches capture the predictability of the IVS of equity options in the IT sector. Both the two-step and one-step approach significantly outperform the benchmark model and the random walk model. This confirms that more sophisticated models are better able to capture the dynamics of the IVS of equity options. The one-step approach is the best performing averaged over time for each individual equity option in terms of performance evaluation measures. This indicates that the IVS prediction accuracy is higher when the factors are estimated simultaneously with the IVS. Second, the consideration of exogenous variables improves the predictions of the approaches in terms of statistical performance. The selection of the exogenous variables is different among the equity options and for each rolling window. The LARS algorithm relatively frequently selects interest rate and expected inflation as variables. Across the equity options, these variables are selected in respectively 39% and 24% of all rolling windows. The (lagged) return of the S&P 500 is found to be a less useful variable than literature suggested. Using the LARS algorithm to select the factors created by PCA and the exogenous variables improves the forecast performance. Third, all the approaches are shown to have economic value. Applying the one-step ahead forecasts of the approaches to an ATM straddle trading strategy generates positive returns before transaction costs are considered. However, when transactions costs are accounted for the profits become negative. From an economic perspective, including exogenous variables to the approaches does not have a large impact. Nonetheless, the predictable dynamics of equity option IVS could be helpful for a trader's prediction and hedging strategies.

However, this study has some limitations regarding the methodology. Since IVS data had an unbalanced panel, the data had to be transformed before predictions could be made. In this paper, one option in each of the 18 maturity-moneyness groups was selected. Considering more groups is possible. Another alternative would be to consider the unbalanced set of all options, since forecasting is most relevant at an individual option level. The performance of the two-step and one-step approach is compared to assess the relevance of simultaneously estimating the factors and predicting the IVS. To effectively compare the approaches, they should differ only on this point. However, as seen in the methodology, the exogenous variables are included in the observation equation for the one-step approach, whereas in the two-step approach they are included in the VAR model for the factors. Additionally, the Kalman filter used in the one-step approach is a recursive algorithm. It therefore uses all the available information up to time t efficiently. In the two-step approach only p lags with a maximum of five are considered for the prediction of the factors. These points need to be kept in mind when assessing the relative performance of the two approaches.

There are three topics that can be addressed in further research. First, the results indicate that the addition of exogenous variables improves forecast accuracy. However, access to daily updates of additional exogenous variables (GDP, unemployment rates, money supply, CPI) was limited. The consideration of these indicators could further increase the forecast accuracy of the IVS of equity options. Second, the economic value of the forecasts is assessed using only a simple trading strategy. Due to some simplifying assumptions the transaction costs when applying the strategy are high. More complex trading strategies could be explored, where for example the amount of trades per day is limited. Third, investigation of combined forecasts of the different approaches for individual equity options is advised to improve the forecast performance. Alternatively, since the one-step approach tends to outperform the other models, a future study could focus on optimising predictions with this approach.

References

J. Bai and S. Ng. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146 (2):304–317, 2008.

E. Bair, T. Hastie, D. Paul, and R. Tibshirani. Prediction by supervised principal components. *Journal of the American Statistical Association*, 101(473):119–137, 2006.

M. Bedendo and S. D. Hodges. The dynamics of the volatility skew: A kalman filter approach. *Journal of Banking & Finance*, 33(6):1156–1165, 2009.

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

N. P. Bollen and R. E. Whaley. Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance*, 59(2):711–753, 2004.

G. Chalamandaris and A. E. Tsekrekos. Can static models predict implied volatility surfaces? evidence from otc currency options. Technical report, Working paper, Athens University of Economics and Business, 2009.

G. Chalamandaris and A. E. Tsekrekos. Predictable dynamics in implied volatility surfaces from otc currency options. *Journal of Banking & Finance*, 34(6):1175–1188, 2010.

N.-F. Chen, R. Roll, and S. A. Ross. Economic forces and the stock market. *Journal of business*, pages 383–403, 1986.

R. Cont, J. Da Fonseca, et al. Dynamics of implied volatility surfaces. *Quantitative finance*, 2(1):45–60, 2002.

F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134–144, 2002.

B. Dumas, J. Fleming, and R. E. Whaley. Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6):2059–2106, 1998.

B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, et al. Least angle regression. *The Annals of statistics*, 32 (2):407–499, 2004.

M. R. Fengler, W. K. Härdle, and E. Mammen. A semiparametric factor model for implied volatility surface dynamics. *Journal of Financial Econometrics*, 5(2):189–218, 2007.

J. R. Franks and E. S. Schwartz. The stochastic behaviour of market variance implied in the prices of index options. *The economic journal*, 101(409):1460–1475, 1991.

S. Goncalves and M. Guidolin. Predictable dynamics in the s&p 500 index options implied volatility surface. *The Journal of Business*, 79(3):1591–1635, 2006.

J. D. Hamilton. State-space models. *Handbook of econometrics*, 4:3039–3080, 1994.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82 (1):35–45, 1960.

B. Kelly and S. Pruitt. The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2):294–316, 2015.

S. Mixon. Factors explaining movements in the implied volatility surface. *Journal of Futures Markets*, 22 (10):915–937, 2002.

R. H. Shumway and D. S. Stoffer. An approach to time series smoothing and forecasting using the em algorithm. *Journal of time series analysis*, 3(4):253–264, 1982.

G. Skiadopoulos, S. Hodges, and L. Clewlow. The dynamics of the s&p 500 implied volatility surface. *Review of derivatives research*, 3(3):263–282, 2000.

J. H. Stock and M. W. Watson. Forecasting with many predictors. *Handbook of economic forecasting*, 1: 515–554, 2006.

J. H. Stock and M. W. Watson. Generalized shrinkage methods for forecasting using many predictors. *Journal of Business & Economic Statistics*, 30(4):481–493, 2012.

M. van der Wel, S. R. Ozturk, and D. van Dijk. Dynamic factor models for the volatility surface. In *Dynamic Factor Models*, pages 127–174. Emerald Group Publishing Limited, 2016.

A Appendix

A.1 Proofs

A general explanation of PCA is given in subsection A.1.1. The LARS algorithm of Efron et al. [2004] is elaborated in subsection A.1.2. A detailed description of the Kalman filter and the expectation maximisation algorithm is found subsection A.1.3.

A.1.1 Principal Compenent analysis

PCA finds the principal components in a dataset. Principal components are the directions where there is most variance in the data. PCA is used for data reduction.

Assume N given data variables $\boldsymbol{x_t} = (x_{1,t}, ..., x_{N,t})'$. Define the factors $f_{i,t}$ as a linear combinations of the variables in the large dataset.

$$\boldsymbol{f}_{i,t} = \boldsymbol{a}_i'\boldsymbol{x}_t = a_{i,1}'x_{1.t} + \cdots + a_{i,N}'x_{N.t} \tag{32}$$

such that,

$$\boldsymbol{f}_t = \boldsymbol{A}'\boldsymbol{x}_t \tag{33}$$

Without loss of generality, $\boldsymbol{A}$ may be taken to be orthogonal, implying that $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}_N$ or $\boldsymbol{A}^{-1} = \boldsymbol{A}'$. Given that $\boldsymbol{f}_t = \boldsymbol{A}'\boldsymbol{x}_t$ and $\boldsymbol{A}$ is orthogonal, it follows that $\boldsymbol{x_t} = \boldsymbol{A}\boldsymbol{f}_t = a_{i,1}f_{1.t} + \cdots + a_{i,N}'f_{N.t}$. The aim is to reduce the dimension to $r < N$ in such a way that the r factors capture the bulk of the variation in the dataset. A principal component analysis identifies the linear combinations of the data that are uncorrelated and have maximum variance. This involves solving the maximisation problem denoted below. Let $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{x}} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{x}_t - \bar{\boldsymbol{x}})(\boldsymbol{x}_t - \bar{\boldsymbol{x}})'$ denote the covariance matrix of the data $\boldsymbol{x}_t$.

$$\max_{\boldsymbol{a}_{1:r}} \quad \boldsymbol{a}_{1:r}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{x}}\boldsymbol{a}_{1:r} \quad \text{subject to} \quad \boldsymbol{a}_{1:r}\boldsymbol{a}_{1:r} = \boldsymbol{I}_r \tag{34}$$

To solve the maximisation problem, form the Lagrangian $L = \boldsymbol{a}_{1:r}'\hat{\boldsymbol{\Sigma}}_{\boldsymbol{x}}\boldsymbol{a}_{1:r} - l(\boldsymbol{a}_{1:r}'\boldsymbol{a}_{1:r} - \boldsymbol{I}_r)$, where l is the Lagrange multiplier. Taking the derivative with respect to $\boldsymbol{a}_{1:r}$ results in $\frac{\delta L}{\delta \boldsymbol{a}_{1:r}} = 2\hat{\boldsymbol{\Sigma}}_{\boldsymbol{x}}\boldsymbol{a}_{1:r} - 2l\boldsymbol{a}_{1:r}$. Setting this equal to zero gives the first order conditions, which can be rewritten as:

$$\boldsymbol{\Sigma}_{IV}\boldsymbol{a}_{1:r} = l\boldsymbol{a}_{1:r} \tag{35}$$

Which shows that the solution $\boldsymbol{a}_{1:r}$ are the r eigenvector of $\hat{\Sigma}_{IV^k}$ according to the definition of an eigenvector.

A.1.2 LARS algorithm

The LARS algorithm of Efron et al. [2004] starts with $\hat{\mu}_0 = 0$ and build the current estimate $\hat{\mu}$ by the steps. The current correlation is defined as $\hat{c} = X'(y - \hat{\mu}_A)$. Define the set K as the indices of the variables which corresponds to the largest absolute correlation

$$K = [j : |\hat{c}_j| = \hat{C}] \qquad \hat{C} = \max_j |\hat{c}_j| \tag{36}$$

Let $s_j = sign\hat{c}_j$ and define the matrix $X_K = (...s_j x_j...)_{j\in K}$. To update the estimates the algorithm tends to find the vector which has equal angles with the existing variables. Define $G_K = X_K' X_K$ and $K_K = (1_K' G_K^{-1} 1_K)^{-1/2}$, where 1_K is a vector of 1's of length equal to the matrix K. The equiangular vector is defined as:

$$G_K = X_K w_K \qquad w_K = K_K G_K^{-1} 1_K \tag{37}$$

This is true because

$$X_K' u_K = K_K 1_K \tag{38}$$

where u_K is a unit vector ($||u_K||^2 = 1$) with columns of the active set, therefore $a_K = X' u_K$. The algorithm updates the estimate in the direction of u_K, so the next step becomes:

$$\hat{\mu}^{new} = \hat{\mu} + \hat{\gamma} u_K \tag{39}$$

where $\gamma > 0$ is the amount the estimate changes with. The updated correlation between the variables is as follows:

$$c_j(\gamma) = x_j'(y - \mu(\gamma)) = x_j'(y - \hat{\mu} - \hat{\gamma} u_K) = \hat{c}_j - \gamma a_j \tag{40}$$

implying that the correlation is decreasing. γ is chosen in order that some other variable has the same absolute correlation with the active set, which lead to values of γ as

$$\hat{\gamma} = \min_{j\in K^c}{}^{+}\Big(\frac{\hat{C} - \hat{c}_j}{A_K - a_j}, \frac{\hat{C} + \hat{c}_j}{A_K + a_j}\Big) \tag{41}$$

The "min^{+}" indicates that only positive values are considered.

A.1.3 Kalman filter

This subsection of the appendix aims to elaborate on the main steps in the Kalman Filter. I use similar notation and equations as Hamilton [1994]. The state space is a representation of a linear time series model given by an equation that describes how the observations are linked to the latent factors, and an equation that describes how the factors evolve over time.

$$\boldsymbol{y}_t = \boldsymbol{H}'\boldsymbol{\xi}_t + \boldsymbol{w}_t \qquad w_t \sim N(0, \boldsymbol{R}) \tag{42}$$

$$\boldsymbol{\xi}_{t+1} = \boldsymbol{F}\boldsymbol{\xi}_t + \boldsymbol{v}_{t+1} \qquad v_{t+1} \sim N(0, \boldsymbol{Q}) \tag{43}$$

for $t = 1, ..., T$, where y_t is the observed $N \times 1$ vector of dependent variables at time t , ξ_t is the (mostly) unobserved $r \times 1$ state vector, $\boldsymbol{H}$ and $\boldsymbol{F}$ are estimated matrices of the appropriate dimensions. The disturbance vectors $\boldsymbol{w}_t$ and $\boldsymbol{v}_{t+1}$ are mutually and serially uncorrelated with zero means and variance matrices

$\boldsymbol{R}$ and $\boldsymbol{Q}$, respectively. The normality condition of the error terms is not necessary. The Kalman filter is a recursive procedure through the time index t that provides an optimal forecast of the states ξ_{t+1} given the information at time t. The parameters $\boldsymbol{H}$, $\boldsymbol{R}$, $\boldsymbol{F}$ and $\boldsymbol{Q}$ are unknown and need to be estimated. The estimation is done by the EM algorithm which first runs the Kalman filter and smoother with intial values, subsequently uses the expected states conditional on the complete information set in the loglikelihood to found the parameters that maximises the loglikelihood function. The EM algorithm iterates these steps until the parameters converge.

The parameters can be estimated by Maximum Likelihood but this is problematic with many states. The Expectation Maximisation (EM) is more robust since the maximisation step is done analytically which is hard for ML. Thereby the EM-algorithm ensures that the covariance matrices $\mathbf{R}$ and $\mathbf{Q}$ are positive definite and symmetric. The EM algorithm takes the conditional expectation of the states conditional on the complete information set and maximises the resulting likelihood function with respect to the unknown parameters.

The aim of the EM algorithm is to maximise the following joint density of observing the states $\boldsymbol{\xi}_{0:T}$ and the data $\boldsymbol{y}_{1:T}$:

$$\begin{aligned} L(\boldsymbol{y}_{1:T}, \boldsymbol{\xi}_{0:T}|\boldsymbol{\theta}) = \frac{T}{2} log|\boldsymbol{R}^{-1}| - \frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t - \boldsymbol{H}'\boldsymbol{\xi}_t)'\boldsymbol{R}^{-1}(\boldsymbol{y}_t - \boldsymbol{H}'\boldsymbol{\xi}_t) \\ + \frac{T}{2} log|\boldsymbol{Q}^{-1}| - \frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{\xi}_t - F\boldsymbol{\xi}_{t-1})'\boldsymbol{Q}^{-1}(\boldsymbol{\xi}_t - F\boldsymbol{\xi}_{t-1}) \end{aligned} \tag{44}$$

The EM algorithm iterates the expectation and maximisation step, which are clarified below, until the parameters are converged.

The **Expectation step** of the EM-algorithm for state space models. The Kalman filter first predicts the states based on the information up to time t in equation 45. Where $\hat{\boldsymbol{\xi}}_{t+1|t}$ represents the states, and $\boldsymbol{P}_{t+1|t}$ the variance of the states. However, when $\boldsymbol{y}_{t+1}$ is observed it is possible to do better than the initial prediction. Therefore the prediction and the variance are updated based on a well-known result for normal variables in equation 46. The prediction $\hat{\boldsymbol{\xi}}_{t+1|t+1}$ is adjusted based on the prediction error, and the variance $\boldsymbol{P}_{t+1|t+1}$ becomes smaller, since there is more information. Given the estimates of the states for the complete sample size T the smoother iterates backwards to compute the smoothed estimates as in equation 47. For the exact derivation of the Kalman filter I refer to Hamilton [1994].

- The Kalman *prediction* step:

$$\hat{\boldsymbol{\xi}}_{t+1|t} = \boldsymbol{F}\hat{\boldsymbol{\xi}}_{t|t}, \qquad \boldsymbol{P}_{t+1|t} = \boldsymbol{F}\boldsymbol{P}_{t|t}\boldsymbol{F}' + \boldsymbol{Q} \tag{45}$$

- The Kalman *updating* step:

$$\begin{aligned}\hat{\boldsymbol{\xi}}_{t+1|t+1} &= \hat{\boldsymbol{\xi}}_{t+1|t} + \boldsymbol{P}_{t+1|t}\boldsymbol{H}(\boldsymbol{H}'\boldsymbol{P}_{t+1|t}\boldsymbol{H} + \boldsymbol{R})^{-1}(\boldsymbol{y}_{t+1} - \boldsymbol{H}'\hat{\boldsymbol{\xi}}_{t+1|t}) \\ \boldsymbol{P}_{t+1|t+1} &= \boldsymbol{P}_{t+1|t} - \boldsymbol{P}_{t+1|t}\boldsymbol{H}(\boldsymbol{H}'\boldsymbol{P}_{t+1|t}\boldsymbol{H} + \boldsymbol{R})^{-1}\boldsymbol{H}'\boldsymbol{P}_{t+1|t}\end{aligned} \tag{46}$$

- The Kalman *smoothing* step:

$$\begin{aligned}\hat{\boldsymbol{\xi}}_{t|T} &= \hat{\boldsymbol{\xi}}_{t|t} + \boldsymbol{P}_{t|t}\boldsymbol{F}'\boldsymbol{P}^{-1}_{t+1|t}(\hat{\boldsymbol{\xi}}_{t+1|T} - \hat{\boldsymbol{\xi}}_{t+1|t}) \\ \boldsymbol{P}_{t|T} &= \boldsymbol{P}_{t|t} - \boldsymbol{P}_{t|t}\boldsymbol{F}'\boldsymbol{P}^{-1}_{t+1|t}(\boldsymbol{P}_{t+1|t} - \boldsymbol{P}_{t+1|T})\boldsymbol{P}^{-1}_{t+1|t}\boldsymbol{F}\boldsymbol{P}_{t|t} \\ \boldsymbol{P}_{t+1,t|T} &= \boldsymbol{P}_{t+1|T}\boldsymbol{P}^{-1}_{t+1|t}\boldsymbol{F}\boldsymbol{P}_{t|t}\end{aligned} \tag{47}$$

The **Maximisation step** of the EM-algorithm for state space models. In this step the output of the Kalman filter and smoother are used as the true values. The maximisation of equation 44 can be done analytically with respect to the parameters $\boldsymbol{\theta} = [\boldsymbol{F}, \boldsymbol{H}, \boldsymbol{Q}, \boldsymbol{R}]$:

$$\begin{aligned}\boldsymbol{F} &= \big(\sum_{t=1}^{T}\hat{\boldsymbol{\xi}}_{t|T}\hat{\boldsymbol{\xi}}'_{t-1|T} + \boldsymbol{P}_{t,t-1|T}\big)\big(\sum_{t=0}^{T-1}\hat{\boldsymbol{\xi}}_{t|T}\hat{\boldsymbol{\xi}}'_{t|T} + \boldsymbol{P}_{t|T}\big)^{-1} \\ \boldsymbol{H}' &= \big(\sum_{t=1}^{T}\boldsymbol{y}_t\hat{\boldsymbol{\xi}}'_{t|T}\big)\big(\sum_{t=1}^{T}\hat{\boldsymbol{\xi}}_{t|T}\hat{\boldsymbol{\xi}}'_{t|T} + \boldsymbol{P}_{t|T}\big)^{-1} \\ \boldsymbol{Q} &= \frac{1}{T}\sum_{t=1}^{T}\hat{\boldsymbol{\xi}}_{t|T}\hat{\boldsymbol{\xi}}'_{t|T} + \boldsymbol{P}_{t|T} - \boldsymbol{F}[\hat{\boldsymbol{\xi}}_{t-1|T}\hat{\boldsymbol{\xi}}'_{t|T} + \boldsymbol{P}_{t-1,t|T}] \\ &\quad - [\hat{\boldsymbol{\xi}}_{t|T}\hat{\boldsymbol{\xi}}'_{t-1|T} + \boldsymbol{P}_{t-1,t|T}]\boldsymbol{F}' + \boldsymbol{F}[\hat{\boldsymbol{\xi}}_{t-1|T}\hat{\boldsymbol{\xi}}'_{t-1|T} + \boldsymbol{P}_{t-1|T}]\boldsymbol{F}'\big) \\ \boldsymbol{R} &= \frac{1}{T}\sum_{t=1}^{T}\big(\boldsymbol{y}_t\boldsymbol{y}'_t - \boldsymbol{H}'\hat{\boldsymbol{\xi}}_{t|T}\boldsymbol{y}'_t - \boldsymbol{y}_t\hat{\boldsymbol{\xi}}'_{t|T}\boldsymbol{H} + \boldsymbol{H}'[\hat{\boldsymbol{\xi}}_{t|T}\hat{\boldsymbol{\xi}}'_{t|T} + \boldsymbol{P}_{t|T}]\boldsymbol{H}\big)\end{aligned} \tag{48}$$

The derivations of the expectation and maximisation step of the EM algorithm can be found in Shumway and Stoffer [1982].

A.2 Data

Table 6 gives the list of considered equity options with the largest market capitalization for the information technology sector in the S&P 500. A glossary and the definitions of the exogenous variables can be found in table 7.

Table 6: List of the ten largest market cap stocks of the S&P500 for the information technology sector

Symbol	**Name**	**Sector**
AAPL	Apple	Information Technology
MSFT	Microsoft	Information Technology
GOOG	Google	Information Technology
INTC	Intel	Information Technology
CSCO	Cisco Systems	Information Technology
ORCL	Oracle	Information Technology
MA	Mastercard	Information Technology
IBM	IBM	Information Technology
NVDA	NVIDIA	Information Technology
ADBE	Adobe Systems	Information Technology

Table 7: Glossary and definitions of the considered exogenous variables.

Symbol	Variable	Definition (source)	Remarks
		Basic series	
r_t^{SPX}	Return of the S&P 500	$1 - \frac{p_t}{p_{t-1}}$ (OptionMetrics)	Percent, Daily, Not Seasonally Adjusted
$E[\pi_t]$	5-Year Breakeven Inflation Rate	Represents a measure of expected inflation derived from 5-Year Treasury Constant Maturity Securities and 5-Year Treasury Inflation-Indexed Constant Maturity Securities (FRED)	Percent, Daily, Not Seasonally Adjusted. The latest value implies what market participants expect inflation to be in the next 5 years, on average.
TB_t	Treasury Bill	The 1-month constant maturity treasury bill rate (FRED)	Percent, Daily, Not Seasonally Adjusted
GB_t	Long government bond	30-Year Treasury Constant Maturity Rate (FRED)	Percent, Daily, Not Seasonally Adjusted
CB_t	High rated corporate bond	Moody's Seasoned Aaa Corporate Bond Yield (FRED)	Percent, Daily, Not Seasonally Adjusted
		Derived series	
I_t	Real interest rate	$TB_t - E[\pi_t]$	
Y_{spread}	Corporate/governement yield spread	$CB_t - GB_t$	

Notes: In the top panel the direct observed variables are elaborated. Reported are the symbol, variable and definition along with the source between brackets and remarks. In the lower panel the derived time series are explained.

Table 8: Correlations of the implied volatilites within the maturity-moneyness groups.

	$\tau < 60$			$60 < \tau < 180$			$\tau > 180$			$\tau < 60$			$60 < \tau < 180$			$\tau > 180$		
	DOTM -call	OTM - call	ATM -call	DOTM -call	OTM - call	ATM -call	DOTM -call	OTM - call	ATM -call	DOTM - put	OTM - put	ATM- put	DOTM - put	OTM - put	ATM - put	DOTM - put	OTM - put	ATM - put
DOTM -call	0.0171	0.9789	0.9796	0.9466	0.9475	0.9505	0.9338	0.9348	0.9378	0.8995	0.9629	0.9778	0.8198	0.9300	0.9481	0.7700	0.9049	0.9346
OTM - call		0.0162	0.9874	0.9513	0.9581	0.9604	0.9432	0.9458	0.9479	0.9078	0.9726	0.9844	0.8186	0.9364	0.9563	0.7682	0.9117	0.9433
ATM -call			0.0167	0.9486	0.9569	0.9625	0.9408	0.9458	0.9498	0.9208	0.9818	0.9908	0.8358	0.9450	0.9610	0.7816	0.9185	0.9469
DOTM -call				0.0135	0.9871	0.9873	0.9822	0.9816	0.9831	0.8553	0.9276	0.9439	0.8311	0.9553	0.9828	0.7997	0.9477	0.9785
OTM - call					0.0128	0.9959	0.9822	0.9896	0.9907	0.8798	0.9353	0.9510	0.8531	0.9672	0.9891	0.8162	0.9515	0.9846
ATM -call						0.0134	0.9825	0.9906	0.9943	0.8913	0.9475	0.9592	0.8685	0.9763	0.9958	0.8308	0.9628	0.9909
DOTM -call							0.0123	0.9887	0.9886	0.8498	0.9204	0.9363	0.8235	0.9522	0.9778	0.7925	0.9440	0.9821
OTM - call								0.0119	0.9964	0.8725	0.9269	0.9408	0.8511	0.9583	0.9849	0.8172	0.9545	0.9883
ATM -call									0.0122	0.8800	0.9350	0.9467	0.8644	0.9697	0.9912	0.8387	0.9668	0.9954
DOTM - put										0.0207	0.9366	0.9265	0.8695	0.8884	0.8943	0.7965	0.8606	0.8801
OTM - put											0.0190	0.9886	0.8570	0.9488	0.9536	0.7960	0.9183	0.9377
ATM - put												0.0174	0.8444	0.9504	0.9623	0.7898	0.9227	0.9473
DOTM - put													0.0203	0.8953	0.8786	0.8497	0.8662	0.8708
OTM - put														0.0159	0.9845	0.8545	0.9658	0.9767
ATM - put															0.0142	0.8406	0.9698	0.9927
DOTM - put																0.0169	0.8954	0.8585
OTM - put																	0.0142	0.9801
ATM- put																		0.0128

Notes: This table contains the correlations and variances of the implied volatilities within the maturity-moneyness groups. The off-diagonal elements corresponds to the correlations and the diagonal elements represent the variances within a group.

A.3 Results

Table 9 contains the ADF and KPSS test statistics and the p-values for unit roots in the exogenous variables. The tables 10 and 11 respectively give the ADF and KPSS test statistics and the corresponding p-values for the factors created by applying PCA to the covariance matrix of the equity options and the deterministic factors. Figure 9 plots the time-series of MAE statistics of the individual equity options for the different approaches. Table 12 and table 13 give the average MAE and MCPDC statistics for the individual equity options for different subperiods. In table 14 and table 15 the performance measures for the individual equity options are found.

Table 9: Unit root test statistics for the exogenous variables used in the different approaches.

	ADF test statistic (p-value)	KPSS test statistic (p-value)
Return S&P500	-54.708 (0.001)	0.0145 (0.100)
Lagged Return S&P500	-54.712 (0.001)	0.0146 (0.100)
Inflation	-4.578 (0.001)	30.671 (0.010)
Interest rate	-19.510 (0.001)	8.945 (0.010)
Yield	-25.848 (0.001)	11.674 (0.010)
S&P 500	-9.247 (0.001)	24.158 (0.010)

Notes: The table contains the test statistics and in brackets the p-values of the ADF test and KPSS test for unit roots in the exogenous variables. The univariate time-series are tested for the complete out-of-sample period.

Table 10: ADF Unit root test statistics for the factors created by PCA, and the deterministic factors used in the benchmark model.

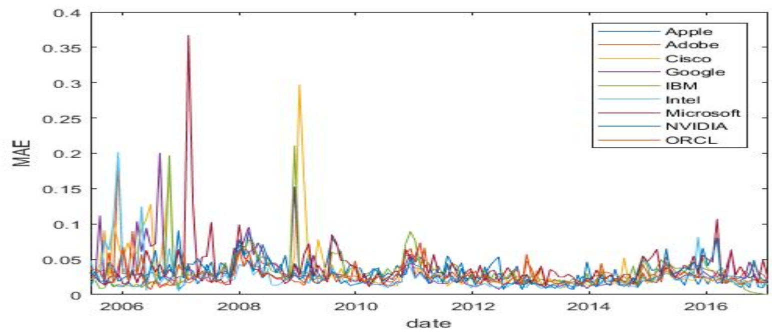
	AAPL	GOOG	MSFT	INTC	CSCO	ORCL	MA	IBM	NVDA	ADBE
PC factor 1	-7.99 (0.001)	-7.22 (0.001)	-6.52 (0.001)	-7.69 (0.001)	-7.48 (0.001)	-7.49 (0.001)	-6.26 (0.001)	-3.18 (0.089)	-4.69 (0.001)	-7.56 (0.001)
PC factor 2	-9.27 (0.001)	-9.47 (0.001)	-12.39 (0.001)	-11.66 (0.001)	-12.26 (0.001)	-10.37 (0.001)	-13.82 (0.001)	-8.79 (0.001)	-10.07 (0.001)	-10.33 (0.001)
PC factor 3	-10.79 (0.001)	-14 (0.001)	-15.66 (0.001)	-18.7 (0.001)	-14.41 (0.001)	-14.57 (0.001)	-25.69 (0.001)	-11.99 (0.001)	-11.98 (0.001)	-11.39 (0.001)
PC factor 4	-12.59 (0.001)	-19.58 (0.001)	-18.12 (0.001)	-20.67 (0.001)	-23.82 (0.001)	-22.79 (0.001)	-27.62 (0.001)	-12.24 (0.001)	-15.76 (0.001)	-15.32 (0.001)
PC factor 5	-18.63 (0.001)	-22.85 (0.001)	-28.81 (0.001)	-22.78 (0.001)	-32.72 (0.001)	-38.69 (0.001)	-15.14 (0.001)	-21.39 (0.001)	-27.68 (0.001)	-25.08 (0.001)
DF factor 1	-9.547 (0.001)	-11.06 (0.001)	-53.26 (0.001)	-46.69 (0.001)	-51.01 (0.001)	-41.91 (0.001)	-11.52 (0.001)	-29.17 (0.001)	-27.53 (0.001)	-50.75 (0.001)
DF factor 2	-26.54 (0.001)	-28.63 (0.001)	-53.24 (0.001)	-50.12 (0.001)	-52.27 (0.001)	-50.8 (0.001)	-30.31 (0.001)	-30.94 (0.001)	-42.71 (0.001)	-50.43 (0.001)
DF factor 3	-14.49 (0.001)	-16.21 (0.001)	-52.91 (0.001)	-31.55 (0.001)	-52.25 (0.001)	-45.43 (0.001)	-22.19 (0.001)	-18.48 (0.001)	-24.59 (0.001)	-27.59 (0.001)
DF factor 4	-11.87 (0.001)	-12.4 (0.001)	-53.15 (0.001)	-47.82 (0.001)	-51.37 (0.001)	-46.89 (0.001)	-17.79 (0.001)	-29.25 (0.001)	-38.15 (0.001)	-50.78 (0.001)
DF factor 5	-26.72 (0.001)	-28.43 (0.001)	-53.14 (0.001)	-50.41 (0.001)	-52.96 (0.001)	-50.76 (0.001)	-30.19 (0.001)	-30.96 (0.001)	-43.25 (0.001)	-50.39 (0.001)

Notes: The table contains the results of the augmented Dickey-Fuller test for unit roots of factors created by PCA, and the deterministic factors used in the benchmark model. The univariate time-series are being for the complete out-of-sample period.

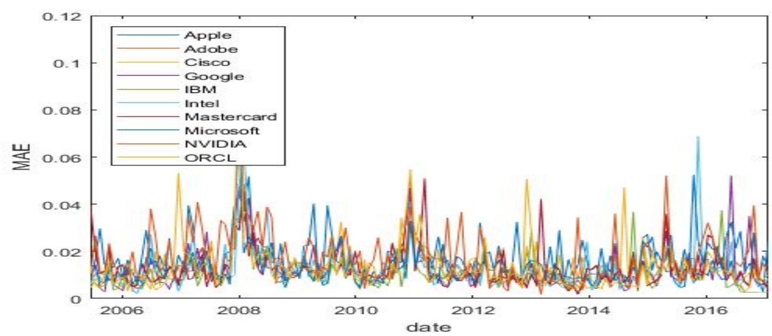
Table 11: KPSS unit root test statistics for the factors created by PCA, and the deterministic factors used in the benchmark model.

	AAPL	MSFT	GOOG	INTC	CSCO	ORCL	MA	IBM	NVDA	ADBE
PC factor 1	2.503 (0.01)	16.137 (0.01)	10.533 (0.01)	6.157 (0.01)	17.098 (0.01)	13.703 (0.01)	27.532 (0.01)	17.933 (0.01)	21.285 (0.01)	7.634 (0.01)
PC factor 2	11.148 (0.01)	3.108 (0.01)	1.345 (0.01)	8.988 (0.01)	7.454 (0.01)	20.954 (0.01)	0.178 (0.02)	18.713 (0.01)	4.385 (0.01)	8.49 (0.01)
PC factor 3	9.895 (0.01)	7.908 (0.01)	1.887 (0.01)	3.089 (0.01)	4.767 (0.01)	0.582 (0.01)	1.481 (0.01)	1.456 (0.01)	19.619 (0.01)	3.164 (0.01)
PC factor 4	0.339 (0.01)	6.549 (0.01)	2.466 (0.01)	14.456 (0.01)	5.151 (0.01)	0.834 (0.01)	0.172 (0.03)	8.292 (0.01)	6.944 (0.01)	3.167 (0.01)
PC factor 5	1.487 (0.01)	6.269 (0.01)	0.682 (0.01)	1.1766 (0.01)	0.282 (0.01)	0.903 (0.01)	6.933 (0.01)	2.375 (0.01)	4.308 (0.01)	5.893 (0.01)
DF factor 1	2.262 (0.01)	0.0236 (0.1)	4.125 (0.01)	0.122 (0.09)	0.021 (0.1)	1.125 (0.01)	22.281 (0.01)	0.246 (0.01)	6.845 (0.01)	0.088 (0.1)
DF factor 2	2.723 (0.01)	0.0615 (0.1)	1.547 (0.01)	0.053 (0.1)	0.015 (0.1)	0.229 (0.01)	1.14 (0.01)	0.07 (0.1)	0.137 (0.07)	0.066 (0.1)
DF factor 3	11.777 (0.01)	0.129 (0.081)	2.397 (0.01)	4.403 (0.01)	0.075 (0.1)	1.513 (0.01)	1.152 (0.01)	8.566 (0.01)	2.721 (0.01)	2.382 (0.01)
DF factor 4	3.933 (0.01)	0.044 (0.1)	0.2371 (0.01)	0.0367 (0.1)	0.042 (0.1)	0.566 (0.01)	2.757 (0.01)	0.108 (0.1)	0.233 (0.01)	0.066 (0.1)

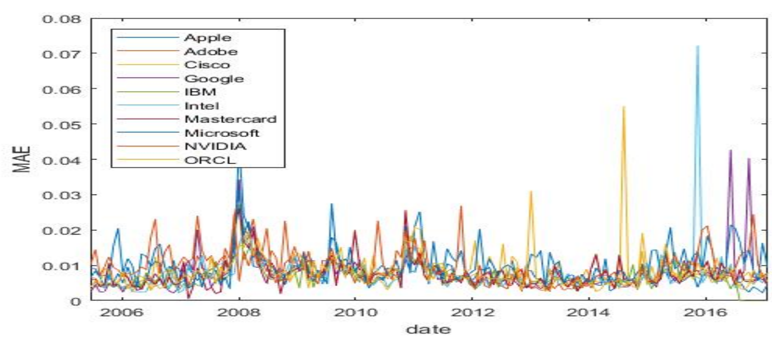
Notes: The table contains the results of Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test for a unit roots of factors created by PCA, and the deterministic factors used in the benchmark model. The univariate time-series are being for the complete out-of-sample period.



(a) Benchmark



(b) Two-step



(c) One-step

Figure 9: Time-series of the MAE statistics for the individual equity options. The upper panel plots the time-series for the benchmark model. The middle panel plots the time-series for the MAE statistics of the forecasts of the two-step approach. The lower panel shows the MAE statistics for the one-step approach.

Table 12: Average MAE statistics

		Before financial crisis *July 18, 2005 until January 31, 2007*	During financial crisis *February 1, 2007 until September 1, 2009*	After financial crisis *September 2, 2009 until December 31, 2017*
Apple				
	benchmark	0.0259	0.0434	0.0272
	Two-step	0.0135	0.0215	0.0132
	One-step	**0.0083**	**0.0114**	**0.0075**
Adobe				
	benchmark	0.0288	0.0367	0.0271
	Two-step	0.0112	0.0195	0.0121
	One-step	**0.0068**	**0.0129**	**0.0075**
Cisco				
	benchmark	0.0826	0.0320	0.0205
	Two-step	0.0127	0.0206	0.0125
	One-step	**0.0068**	**0.0113**	**0.0075**
Google				
	benchmark	0.0544	0.0485	0.0200
	Two-step	0.0109	0.0221	0.0108
	One-step	**0.0060**	**0.0125**	**0.0070**
IBM				
	benchmark	0.0491	0.0429	0.0288
	Two-step	0.0089	0.0171	0.0116
	One-step	**0.0055**	**0.0094**	**0.0063**
Intel				
	benchmark	0.0549	0.1678	0.0190
	Two-step	0.0110	0.0212	0.0109
	One-step	**0.0059**	**0.0114**	**0.0069**
Mastercard				
	benchmark	0.0344	0.0394	0.0326
	Two-step	0.0208	0.0263	0.0217
	One-step	**0.0164**	**0.0194**	**0.0172**
Microsoft				
	benchmark	0.0357	0.0718	0.0331
	Two-step	0.0154	0.0213	0.0152
	One-step	**0.0103**	**0.0117**	**0.0098**
NVIDIA				
	benchmark	0.0280	0.0497	0.0274
	Two-step	0.0164	0.0273	0.0148
	One-step	**0.0100**	**0.0135**	**0.0085**
Oralce				
	benchmark	0.0324	0.0320	0.0242
	Two-step	0.0121	0.0206	0.0122
	One-step	**0.0072**	**0.0113**	**0.0079**

Notes: The table contains the average MAE statistics of the one-step ahead prediction for the individual equity options. For each equity option the MAE statistic for the three approaches is reported. The statistics are averaged over different subperiods: before, during and after the financial crisis. Bold values represents the lowest MAE statistic for the equity option in the specific sub-period.

Table 13: Average MCPDC statistics

		Before financial crisis *July 18, 2005 until* *January 31, 2007*	During financial crisis *February 1, 2007 until* *September 1, 2009*	After financial crisis *September 2, 2009 until* *December 31, 2017*
Apple				
	benchmark	0.5205	0.5097	0.5157
	Two-step	0.5242	0.5300	0.5367
	One-step	**0.7199**	**0.7832**	**0.7278**
Adobe				
	benchmark	0.5275	0.5229	0.5119
	Two-step	0.5682	0.5675	0.5526
	One-step	**0.7262**	**0.7373**	**0.7100**
Cisco				
	benchmark	0.5310	0.5323	0.5312
	Two-step	0.5603	0.5499	0.5597
	One-step	**0.7617**	**0.7720**	**0.7171**
Google				
	benchmark	0.5063	0.5038	0.5357
	Two-step	0.5676	0.5090	0.5698
	One-step	**0.7573**	**0.7437**	**0.7266**
IBM				
	benchmark	0.5147	0.5143	0.4694
	Two-step	0.5552	0.5508	0.4981
	One-step	**0.6924**	**0.7582**	**0.7152**
Intel				
	benchmark	0.5196	0.5298	0.5414
	Two-step	0.5848	0.5477	0.5705
	One-step	**0.7775**	**0.7665**	**0.7316**
Mastercard				
	benchmark	0.5372	0.5241	0.5216
	Two-step	0.5777	0.5724	0.5622
	One-step	**0.7310**	**0.7469**	**0.7148**
Microsoft				
	benchmark	0.5033	0.5202	0.5135
	Two-step	0.5177	0.5557	0.5406
	One-step	**0.6827**	**0.7739**	**0.6864**
NVIDIA				
	benchmark	0.5234	0.5334	0.5304
	Two-step	0.5609	0.5625	0.5572
	One-step	**0.7467**	**0.7751**	**0.7495**
Oralce				
	benchmark	0.5388	0.5323	0.5324
	Two-step	0.5894	0.5499	0.5553
	One-step	**0.7512**	**0.7720**	**0.7172**

Notes: The table contains the average MCPDC statistics of the one-step ahead prediction for the individual equity options. For each equity option the MCPDC statistic for the three approaches is reported. The statistics are averaged over different subperiods: before, during and after the financial crisis. Bold values represents the lowest MCPDC statistic for the equity option in the specific sub-period.

Table 14: Forecast performance measures for the individual equity options.

	RMSE			MAE			MCPDC		
	h = 1	h = 5	h = 5	h = 1	h = 5	h = 5	h = 1	h = 5	h = 5
Benchmark model									
Apple	0.037	0.037	0.037	0.027	0.027	0.028	0.517	0.518	0.517
Adobe	0.036	0.037	0.038	0.028	0.028	0.032	0.516	0.517	0.517
Cisco	0.043	0.044	0.044	0.036	0.036	0.036	0.531	0.528	0.528
Google	0.035	0.035	0.035	0.028	0.031	0.030	0.529	0.529	0.531
IBM	0.046	0.044	0.045	0.034	0.032	0.033	0.480	0.481	0.481
Intel	0.035	0.034	**0.032**	0.028	0.026	**0.025**	**0.536**	**0.536**	**0.538**
Mastercard	0.045	0.046	0.047	0.033	0.034	0.037	0.525	0.526	0.526
Microsoft	0.052	0.052	0.053	0.034	0.034	0.034	0.511	0.511	0.513
NVIDIA	0.035	0.036	0.035	0.028	0.028	0.028	0.529	0.528	0.528
Oracle	**0.034**	**0.034**	0.033	**0.026**	**0.026**	0.026	0.534	0.535	0.535
Two-step approach									
Apple	0.017	0.019	0.020	0.013	0.014	0.015	0.534	0.534	0.531
Adobe	0.016	0.016	0.017	0.012	0.013	0.014	0.556	0.552	0.551
Cisco	0.017	0.017	0.018	0.013	0.013	0.014	0.560	0.557	0.551
Google	0.014	**0.015**	**0.015**	**0.011**	0.011	**0.012**	0.569	0.568	0.567
IBM	**0.014**	0.906	0.016	0.011	0.745	0.013	0.512	0.511	0.511
Intel	0.014	0.015	0.016	0.011	**0.011**	0.012	**0.574**	**0.570**	**0.569**
Mastercard	0.022	0.023	0.024	0.021	0.022	0.023	0.566	0.562	0.560
Microsoft	0.023	0.024	0.026	0.015	0.016	0.018	0.535	0.532	0.529
NVIDIA	0.020	0.021	0.022	0.015	0.016	0.017	0.558	0.556	0.549
Oracle	0.016	0.017	0.017	0.012	0.013	0.013	0.564	0.561	0.558
One-step approach									
Apple	0.010	0.013	0.016	0.008	0.010	0.012	0.726	0.693	0.667
Adobe	0.010	0.012	0.014	0.007	0.009	0.011	0.714	0.683	0.656
Cisco	0.010	0.012	0.015	0.007	0.009	0.011	0.728	0.683	0.656
Google	0.009	0.011	0.013	0.007	0.008	0.010	0.734	0.697	0.673
IBM	**0.008**	**0.009**	**0.012**	**0.006**	**0.007**	**0.009**	0.710	0.682	0.658
Intel	0.009	0.011	0.013	0.007	0.008	0.010	0.743	0.700	0.669
Mastercard	0.011	0.013	0.015	0.017	0.019	0.021	0.719	0.688	0.661
Microsoft	0.014	0.019	0.023	0.010	0.013	0.015	0.685	0.653	0.634
NVIDIA	0.012	0.013	0.016	0.009	0.010	0.012	**0.749**	**0.721**	**0.696**
Oracle	0.011	0.012	0.013	0.008	0.009	0.010	0.725	0.700	0.676

Notes: The table contains the average root mean squared error (RMSE), average mean absolute error (MAE) and the average mean correct prediction of the direction of change in implied volatilities (MCPDC) across the out-of-sample time period. The performance is evaluated for three different forecast horizons: 1 day, 3 days and 5 days ahead. The different factors sets are indicated as follows: Benchmark indicates the deterministic factors set (benchmark model), Two-step represents the two-step approach where the factors are created by PCA and One-step indicates the one-step approach where a state space model for the implied volatility is estimated using the Kalman filter. The bold values indicate the best performing model within each maturity category.

Table 15: Forecast performance measures for the individual equity options.

	RMSE			MAE			MCPDC		
	h = 1	h = 5	h = 5	h = 1	h = 5	h = 5	h = 1	h = 5	h = 5
Benchmark model including exogenous variables									
Apple	0.037	0.038	0.039	0.027	0.028	0.029	0.519	0.520	0.520
Adobe	0.037	0.039	0.041	0.029	0.031	0.033	0.518	0.518	0.521
Cisco	0.051	0.057	0.059	0.042	0.047	0.047	0.525	0.529	0.528
Google	0.038	0.039	0.040	0.030	0.032	0.035	0.529	0.530	0.532
IBM	0.045	0.047	0.047	0.033	0.034	0.036	0.481	0.481	0.484
Intel	**0.032**	**0.029**	**0.030**	**0.025**	**0.023**	**0.024**	**0.540**	**0.542**	**0.541**
Mastercard	0.045	0.047	0.049	0.038	0.040	0.042	0.525	0.525	0.528
Microsoft	0.052	0.053	0.054	0.034	0.035	0.036	0.514	0.516	0.517
NVIDIA	0.034	0.036	0.037	0.027	0.028	0.029	0.529	0.530	0.529
Oracle	0.035	0.035	0.037	0.026	0.027	0.029	0.538	0.538	0.540
Two-step approach including exogenous variables									
Apple	0.018	0.020	0.022	0.013	0.015	0.017	0.531	0.530	0.528
Adobe	0.016	0.019	0.021	0.012	0.015	0.017	0.555	0.551	0.548
Cisco	0.017	0.019	0.021	0.013	0.015	0.017	0.562	0.556	0.552
Google	**0.015**	**0.017**	**0.019**	**0.012**	**0.014**	**0.016**	0.566	0.562	0.560
IBM	0.016	0.019	0.022	0.013	0.016	0.018	0.516	0.513	0.511
Intel	0.015	0.017	0.019	0.012	0.014	0.016	**0.570**	**0.567**	**0.565**
Mastercard	0.018	0.020	0.023	0.020	0.023	0.025	0.556	0.552	0.549
Microsoft	0.023	0.026	0.029	0.016	0.018	0.021	0.536	0.534	0.532
NVIDIA	0.020	0.023	0.025	0.015	0.018	0.020	0.554	0.547	0.543
Oracle	0.017	0.020	0.022	0.013	0.016	0.018	0.562	0.558	0.555
One-step approach including exogenous variables									
Apple	**0.006**	0.011	0.015	0.005	0.008	0.011	0.791	0.733	0.696
Adobe	0.007	0.011	0.014	0.006	0.008	0.011	0.752	0.699	0.664
Cisco	0.008	0.012	0.015	0.006	0.008	0.011	0.771	0.710	0.677
Google	0.007	0.009	0.012	**0.005**	0.007	0.009	0.778	0.727	0.691
IBM	0.006	0.009	**0.012**	0.005	**0.007**	**0.009**	0.713	0.676	0.646
Intel	0.007	**0.010**	0.012	0.005	0.008	0.009	0.776	0.721	0.690
Mastercard	0.012	0.015	0.019	0.015	0.018	0.020	0.760	0.708	0.672
Microsoft	0.011	0.019	0.024	0.008	0.012	0.014	0.726	0.684	0.661
NVIDIA	0.009	0.013	0.016	0.006	0.009	0.011	**0.805**	**0.755**	**0.719**
Oracle	0.008	0.011	0.013	0.006	0.008	0.009	0.768	0.723	0.691

Notes: The table contains the average root mean squared error (RMSE), average mean absolute error (MAE) and the average mean correct prediction of the direction of change in implied volatilities (MCPDC) across the out-of-sample time period. The performance is evaluated for three different forecast horizons: 1 day, 3 days and 5 days ahead. The different factors sets are indicated as follows: Benchmark indicates the deterministic factors set (benchmark model), Two-step represents the two-step approach where the factors are created by PCA and One-step indicates the one-step approach where a state space model for the implied volatility is estimated using the Kalman filter. The bold values indicate the best performing model within each maturity category.