Erasmus University Rotterdam

Erasmus School of Economics

Bachelor's Thesis

ECONOMETRICS & OPERATIONS RESEARCH

Forecasting the implied volatility surface dynamics of equity options

---

Supervisor: X. Gong

Second assessor: D.J.C. van Dijk

Author: P. Zomerdijk

Student number: 400974

Date: July 2, 2017

Abstract

In this paper we examine the predictability of implied volatility surface dynamics of equity options. In particular, we are focused on studying the predictive performances of models that include implied volatility surface dynamics of S&P500 index options and historical VIX Term Structure information. We find that models incorporating these variables in the form of exogenous autoregressive elements are outperformed in terms of prediction error and forecast accuracy measures by more parsimonious models—such as random walk.

Keywords: Equity options, Index options, implied volatility surface, dynamics, VIX Term Structure, predictability

Contents

List of Figures

List of Tables

1 Introduction

Volatility has always played a central part in financial markets, from the perspectives of traders and risk managers to regulators. Correct specifications and predictions of volatility are likely to improve returns and accurately quantify downside risks[(1)]. Especially when it comes to option pricing (models), volatility is of fundamental value. Nonetheless, volatility is not observable and needs to be estimated. Various measures for volatility exist, one of which is *implied volatility*. Whereas most measures of volatility are based on historic data, implied volatility is forward-looking. Implied volatilities are typically obtained by solving the Black-Scholes model for the unknown volatility parameter, given the quoted option price in the market, underlying asset price and other known variables. However, contrary to Black and Scholes' (1973) option pricing model, volatilities implied from market prices are not constant across strike prices (referred to as the *volatility smile*) and time-to-maturity (referred to as the *term-structure*), together forming the *implied volatility surface (IVS)*. The volatility smile is the famous phenomenon of excessive Black-Scholes implied volatilities for options on a given expiration whose strike prices differ substantially from the current underlying price. Canina and Figlewski (1989) and Rubinstein (1994) provide examples of volatility smiles by plotting implied volatilities against moneyness. Campa and Chang (1995) show that implied volatilities are also related to the time-to-maturity space. This is called the term-structure of implied volatility and refers to the different slope and curvature in shape across maturity given a strike price.

Shortly after the Black-Scholes model was published, Black (1976) already noted that implied volatility might be a function of the underlying price, questioning the constant volatility assumption. Yet, before the market crash of 1987 (Black Monday), equity options trading on American markets did not show a volatility smile. This empirical change is often ascribed to an increase in the probability assessment of extreme (downward) returns by market participants. Whereas the original Black-Scholes-Merton model assumes that stock prices follow a geometric Brownian motion with constant drift and volatility, subsequent historic events have repeatedly led to growing, contradicting evidence (Russian Financial Crisis/LTCM collapse, Dotcom-Bubble, 2007-2009 Global Financial Crisis, October 27 mini-crash and 2010 Flash Crash) of 'extremes beyond the bell curve'. Although some events are characterized by a high degree of—as Nassim Nicholas Taleb would describe in *The Black Swan*—retrospective predictability (Dotcom-Bubble, 2008 Global Financial Crisis), others have a less self-evident origin. In a paper published in 1996, Jens Carsten and Mark Rubinstein calculate that Black Monday was a 27-standard-deviation event; extremely unlikely to occur in a randomized, Brownian motion world. Such evidence indicates negatively skewed, excessive kurtosis return distributions. While one part of the existing literature focuses on finding the correct distributional assumption, another part of the literature is focused on modeling the *IVS* (dynamics) of individual equity or index options.

Unlike estimates based on historical data, in an efficient market, *IVS* predictions do not depend on historical prices or volatilities, as they adapt instantaneously to new stock price realizations, which reflect all (publicly) known information (Sun and Ji, 2015). "Hence, acknowledging the limitations of the Black-Scholes model, traders keep having to change the volatility assumption in order to match market prices" (Bloch, 2012). Therefore, *IV* surfaces represent current market beliefs of risk (Bakshi et al., 2000) and contain (forward-looking) information on the asset price process and its dynamics. There exists a clear relevance for market participants to produce reliable *IVS* predictions,

[(1)] *It's a wrong perception to believe that you can eliminate risk just because you can measure it.* - Professor Robert C. Merton, Nobel Laureate in Economic Sciences, 1997 as "it provides an up-to-date indication of where the market stands with reference to each specific underlying assets" (Bernales and Guidolin, 2014), which evidently can be beneficial for trading strategies. This gives rise to financial engineers trying to exploit *IVS* predictability. Dumas et al. (1998), Heston and Nandi (2000) and Goncalves and Guidolin (2006) find that *IV* surfaces can be successfully modeled.

It is a well known supposition that strong links exist between the volatility of individual equity returns and market returns (think of the Capital Asset Pricing Model, in which the β-coefficient reflects asset riskiness as a function of volatility). However, a similar relationship regarding implied volatilities has received far less attention. Previous literature has mainly focused on the predictability of index option *IV* surfaces. Goncalves and Guidolin (2006) and Bernales and Guidolin (2014) are one of the first with a focus on the potential existence of a dynamic relationship between equity and index *IV* surfaces. Despite alternative ways (Burke, 1988; Heston's SV model, 1993; and Borovkova and Permana, 2009), Bernales and Guidolin (2014) use the framework of Dumas et al. (1998) to model implied volatilities—as a function of the time-to-maturity and (a transformation of) strike price. According to Goncalves and Guidolin (2006), the best fit is achieved by this deterministic model described by Dumas et al. (1998). Goncalves and Guidolin (2006) and Bernales and Guidolin (2014) proceed by modeling the dynamic relation between equity and index *IVS* parameter time series with the use of vector autoregression (VAR) models. They find evidence of strong cross-sectional linkages between the dynamics of the *IV* surface of equity and S&P500 index options. However, despite a good fit, profits vanish when transaction costs are taken into account. The latter has been pointed out by Figlewski (1989) as well. Another approach is recently investigated by Christoffersen et al. (2016), who apply principal component analysis to model the cross-sectional variation in the Dow Jones Index.

In this paper we investigate the joint dynamics of equity and index *IV* surfaces. In particular we try to find evidence for any gain from including information on S&P500 index option (symbol: SPX) *IVS* dynamics in modeling equity option *IVS* dynamics. Intuitively, the modest trading frequency in a large number of equity option contracts might cause slower incorporation of information into prices (and thus *IVS*). Therefore, due to the highly liquid market in index options, there might arise an information asymmetry between equity and SPX *IVS*. If equity and index *IV* surfaces are dynamically related, one might exploit the SPX *IVS* to predict equity *IVS*.

Bernales and Guidolin (2014) use a deterministic model to fit the index and equity *IVS* separately. They subsequently model the dynamics of individual equity *IV* surfaces using a dynamic, exogenous VAR (VARX) model, which includes lags of the equity *IVS* parameters and lags of the SPX *IVS* parameters. We use a similar approach. Additionally, we extend the existing literature by proposing a different method of incorporating information from the dynamics of the S&P500 index option *IVS* in the *IV* surface of equity options. As opposed to Goncalves and Guidolin (2006) and Bernales and Guidolin (2014), this method does not require the estimation of a deterministic *IVS* model for SPX. Instead, it includes information on the expectations of market volatility at key points along the SPX implied volatility term structure directly, using the VXST (9-day), VIX (30-day), VXV (3-month) and VXMT (6-month) volatility indexes of the Chicago Board Options Exchange (CBOE). To the best of our knowledge, there is no existing literature on this topic yet. The main reason for this might be the fact that most of these indexes are relatively new (especially VXST) and therefore the sets of historical observations are (but becoming less) scarce. The VIX dates from 1990, but the VXV, VXMT and VXST only date from 2007, 2008 and 2011, respectively.

The performance of the VARX model including VIX Term Structure data is in this research proven to be fairly equal (if not better) to the dynamic equity-SPX *IVS* model of Bernales and Guidolin (2014). However, we find that the models incorporating exogenous autoregressive dynamics (such as SPX *IVS* dynamics or VIX Term Structure data) are outperformed in terms of prediction error and forecast accuracy measures by more parsimonious, static models (such as a Random Walk model).

This paper has the following setup. The first part of the paper focuses on the data preparation process. Next, we elaborate on the econometric methods used and present estimation results and in-sample fit. In the subsequent section we explain the forecasting procedure and analyze forecasting performance (in terms of implied volatilities and option prices) compared to benchmark models according to basic economic criteria—such as root mean squared prediction error, mean absolute prediction error and, most interestingly, mean correct prediction of direction of change.

2 Data

This research makes use of an extensive amount of data, spanning the pre-financial crisis time period from January 4, 1996 to December 29, 2006 (2768 trading days), for S&P500 index options (symbol: SPX) as well as the options on a total of 100 stocks. The choice of stocks is determined as follows. From the set of all historic S&P500 constituents (813), we remove all stocks that were not continuously part of the index throughout the sample period. Similar to Bernales and Guidolin (2014), we extract from the remaining set of stocks (262) the 150 stocks with highest average daily trading volume. From this set we randomly select 100 stocks to include in our research. A list of the relevant stocks can be found in Appendix B. Information on the S&P500 constituents is available from *Compustat*, through the *Wharton Research Data Services* (WRDS). The option data consists of a Pricing Date, Expiration Date, Strike Price, Daily best closing bid and ask price, Daily traded Volume and Implied Volatility for all (put and call) options (on the S&P500 index and the selection of stocks) traded on all US option trading venues. This data is downloaded from the *OptionMetrics* database, to which WRDS provides access as well. In terms of maturity, between 3 to 6 different index contracts and between 1 to 4 different equity contracts are traded each day. Given some maturity, contracts with different strike prices are available. SPX options and equity options have different exercise styles; American-style (exercisable at any time until expiration) and European-style (only exercisable at expiration), respectively. Additionally, we collect the daily closing prices and declared dividends (announcement date, amount, ex-dividend date, payment date) on the underlying stocks (source: *Compustat*). Furthermore, we extract the daily US Treasury spot rates (1 month, 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, 7 years, 10 years, 20 years, 30 years)[(2)] using *Quandl*'s API for Python.

The cleaning process of the option data is largely based on the criteria applied by Bakshi et al. (1997). First of all, equity contracts with a (mid) price less than \$0.30 and index contracts with a (mid) price less than \$6/16 are removed. The mid price is defined as the midpoint of (highest) bid and (lowest) ask quote. Prices quoted lower than the specified limits are considered to be in (too) close proximity to the minimum tick size of \$0.05 and \$1/16 for equity and index options, respectively. Guidolin and Goncalves (2006) and Bernales and Guidolin (2014) use this similar criterion in order to reduce the influence of price discreteness. Second, to mitigate liquidity-related biases, we remove contracts with fewer than 6 or more than 366 calendar days to expiration. Third, following the procedures of Dumas et al. (1998) and Heston and Nandi (2000), we exclude all contracts with a moneyness outside the [0.9, 1.1] interval. We define the moneyness of a contract with strike price K and underlying closing price S as K/S. We do not remove contracts with zero traded volume, because the quotes still provide useful information and represent trading opportunities, neither do we want to further reduce the already parsimonious set of daily equity contract observations. Finally, contracts that fall outside the following basic no-arbitrage conditions are excluded as well (Sun and Ji, 2015):

American exercise

$$\begin{array}{ll} \text{Call} & max(0, S_t^j - K_i) \leq C(j, K_i, \tau_{it}) \leq S_t^j \\ \text{Put} & max(0, K_i - S_t^j) \leq P(j, K_i, \tau_{it}) \leq K_{it} \end{array}$$

European exercise

$$\begin{array}{ll} \text{call} & max(0, S_t^j - PV(K_i)) \leq c(j, K_i, \tau_{it}) \leq S_t^j \\ \text{put} & max(0, PV(K_i) - S_t^j) \leq p(j, K_i, \tau_{it}) \leq PV(K_{it}) \end{array}$$

(2) One-month spot rates are missing in the beginning of the sample period and are in that case set equal to the 3-month spot rates.

Dividends are not taken into account in the above basic lower and upper no-arbitrage bounds to save computation time. $C/c(P/p)(j, K_i, \tau_{it})$ is the current value of an American/European call (put) option characterized by strike price i, underlying j and time-to-maturity τ at t. S_t^j is the closing price at t of underlying asset j. The present value (PV) of a contract's strike price (K), PV(K), is calculated by multiplying K by the contract's discount factor (see *Methodology* section). We use the conventional notation of denoting American-style options with capital letters and European-style options with lowercase letters.

To provide some insight in the option data that remains after cleaning, we present some summary statistics in Table 1. The data is partitioned in 15 subcategories across moneyness (rows) and maturity (columns), and equity and index option data is analyzed separately.

Table 1. Summary statistics for cleaned data from January 4, 1996 to December 29, 2006

	Short-term (6 < calendar days ≤ 120)			Medium-term (120 < calendar days ≤ 240)			Long-term (240 < calendar days ≤ 366)		
	freq.	mean IV	std. IV	freq.	mean IV	std. IV	freq.	mean IV	std. IV
Equity Options									
M < 0.94	73.12	32.80	9.05	57.78	30.24	7.46	53.54	29.03	6.77
0.94 < M ≤ 0.98	83.73	30.84	8.76	62.91	29.65	7.43	56.51	28.59	6.81
0.98 < M ≤ 1.02	89.02	29.70	8.79	68.68	29.16	7.52	59.80	28.18	6.87
1.02 < M ≤ 1.06	87.64	29.44	8.66	72.15	28.74	7.52	62.08	27.73	6.88
1.06 < M	76.64	30.52	8.68	71.89	28.49	7.61	62.80	27.61	6.93
S&P500 Index Options									
M < 0.94	100	24.71	6.82	94.66	22.29	5.01	70.52	22.31	4.89
0.94 < M ≤ 0.98	100	20.61	5.99	97.21	20.38	4.90	78.15	20.72	4.83
0.98 < M ≤ 1.02	100	17.75	5.82	98.94	19.12	4.85	88.76	19.70	4.77
1.02 < M ≤ 1.06	100	16.67	5.82	94.17	18.11	4.70	74.20	18.99	4.63
1.06 < M	99.13	18.38	6.64	89.80	17.38	4.63	69.93	18.20	4.63

Table 1: Summary of statistics for the data categorized across moneyness (5 subgroups, rows) and maturity (3 subgroups, columns). All values are in terms of percentages. All values are calculated on the data remaining after applying the filtering procedures. The *freq.* column denotes the average trading frequency within the corresponding category (row, column). The trading frequency of (the options on) a particular underlying is defined as the percentage of days on which at least one of all quoted options on that underlying has non-zero traded volume. Note that all values in the *Equity Options* section of the table are averages across all 100 underlying assets and across all days in the sample. The values in the *S&P500 Index Options* section are averages across all days in the sample.

The table clearly reveals some patterns. For example, in case of equity options, average implied volatility consistently decreases across all moneyness subcategories as maturity increases. This is most likely caused because options tend to become less liquid (bigger bid-ask spread, lower trading volume) as maturity increases. This trend is visible by looking at the trading frequency columns in the table, and is true for equity options as well as S&P500 index options. At the same time, the difference in (average) trading frequency between equity and index options is remarkable, but not unexpected—the market in S&P500 options is considered the most liquid option market in the US. Bernales and Guidolin (2014) point out that this difference in trading frequency might imply that the speed at which changes in the shape of the IV surfaces are incorporated into prices is higher for S&P500 options than for equity options. If true, and if the IVS of S&P500 and equity options is related, then the IVS of the S&P500 options could be used to form beliefs about the future shape of equity IV surfaces. To provide illustrative evidence for a potential relationship between the dynamics of index and equity IV surfaces, a few fitted surfaces are shown in Figure 1.

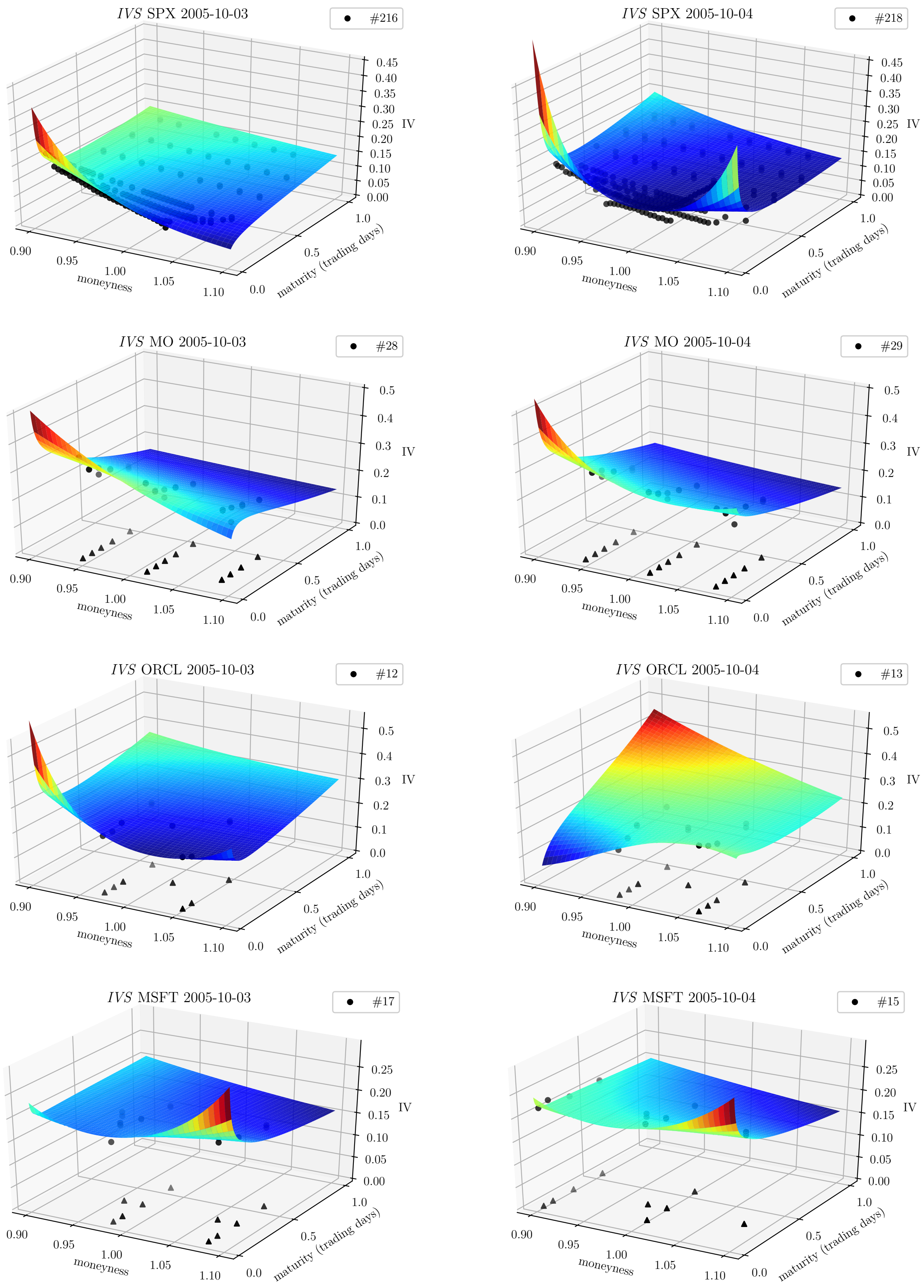

Figure 1: Each figure corresponds to a particular date and underlying (stock or index). All observed, tradeable options (remaining after data cleaning) are plotted as circle markers in the implied volatility space corresponding to their moneyness and maturity coordinates (only in the case of equity options: denoted by triangle markers on the grid). The color-mapped surface illustrates the fitted deterministic *IVS* model (see *Methodology* Section) and should be a proxy for the market implied volatility across the moneyness and maturity grid (0.9 to 1.1 moneyness, 7/366 to 366/366 maturity in years). The legend at the top right displays the number of observed option contracts on the corresponding underlying and date (after data cleaning). The z-axis (IV) denotes the implied volatility (%) in decimal notation. Company names corresponding to tickers can be looked up in Appendix B.

From Figure 1, note the immense difference in the daily number of tradeable options (within our filtering criteria) between the index (SPX) and individual equities (MO, ORCL and MSFT). The shape of the fitted surface is, especially in the case of a scarce set of contracts, highly sensitive to changes in contract parameters and the observations included/excluded. This is explained because in practice only a few market prices (and thus Black-Scholes implied volatilities) are observed for options, at certain strikes and maturities (circle markers in Figure 1). Therefore, the market is incomplete and numerous IV surfaces may be fitted through the available data. This may lead to seemingly large changes in IVS shapes from one day to another (see ORCL). It should be emphasized though, that every surface represents the Ordinary Least Squares (OLS) solution through all observed contracts. Areas of the surface with no (near) observations are indubitably less reliable as a proxy for market implied volatility. Despite this fact, the sub-figures do appear to show a relationship to one another cross-sectionally. In case of MO, the IVS on October 3, 2005 displays similar characteristics as the IVS of SPX on the same day. Also, the IVS of MO and SPX seems to have evolved in sync towards the end of the consecutive day (October 4, 2010). In case of ORCL and MSFT such a relationship is not easily visible. Evidence for a dynamic relationship (between lagged SPX and current equity IV surfaces) is not self-evident from looking at the figure. This will be the focus of the following parts of this paper.

3 Methodology

The first step in investigating the cross-sectional dynamics of the IV surface of equity options is by recursively fitting a deterministic IVS model to options data on the same underlying asset, at daily frequency. This yields 101 (1 index plus 100 stocks) time series of daily (estimated) coefficients. The fitted IV surfaces in Figure 1 are constructed using these coefficients. We then model the time series of coefficients associated with each individual underlying, by fitting an exogenous VAR (VARX) model. The VARX model expresses the deterministic IVS coefficients at time t as a function of its lag(s) and lag(s) of the deterministic IVS coefficients associated with the index.

3.1 Implied Volatility Surface

Similar to Goncalves and Guidolin (2006) and Bernales and Guidolin (2014), we use the following deterministic model to extract the IVS:

$$ln\sigma_{jit} = \beta_{0,it} + \beta_{1,it}M_{jit} + \beta_{2,it}M_{jit}^2 + \beta_{3,it}\tau_{jit} + \beta_{4,it}(M_{jit}\tau_{jit}) + \epsilon_{jit} \tag{1}$$

where i is the i-th underlying in our sample, $i = 1, ..., 101$ (100 equities and the S&P500), j is the j-th tradeable option contract, $j = 1, ..., J$ on trading day t (calculated as the number of calendar days to expiration divided by 366), with $t \in$[01/04/1996 (1), 12/29/2006 (2768)]. The implied volatility of contract j on underlying i on day t is given by σ_{jit}. Likewise, τ_{jit} denotes time to maturity of contract j on underlying i on day t and M_{jit} denotes the time-adjusted moneyness of contract j on underlying i on day t (Eq. (2)). The advantage of modeling the log of implied volatility is that it can not produce negative values. Another advantage of the above model is the fact that all variables are observable. Model (1) is estimated for each individual underlying, on each day, using all observed option contracts on the corresponding underlying, on the corresponding day. This yields a time series with daily estimated coefficients (of size 2768 by 5) for each unique underlying (101 in total).

In the existing literature there exists a broad range of ways to specify (time-adjusted) moneyness. For example, Tompkins (2001) and Tompkins and D'Ecclesia (2006) express the strike price in terms of standard deviation. Goncalves and Guidolin (2006) use a slightly simplified version—without dividends. A few years later Bernales and Guidolin (2014) add dividends to the specification proposed by Goncalves and Guidolin (2006). We follow their definition:

$$M_{jit} = \frac{\log\left(\dfrac{K_{jit}}{\exp(1/d_{jit}^{T=\tau_{jit}} - 1)S_i - FVD_{jit}}\right)}{\sqrt{\tau_{jit}}} \tag{2}$$

where for some trading day t and some option j with underlying asset i, K_{jit} is the corresponding strike price, $d^{T=\tau_{jit}}$ the discount factor at day t with T discounting days (Eq.(3)), S_i the corresponding underlying closing price, τ_{jit} the corresponding time to maturity (in fraction of years) and FVD_{jit} the future value of all announced, forthcoming dividends (Eq.(4)). Out-of-the-money call and in-the-money put options will have a positive M_{jit} and in-the-money call and out-of-the-money puts will have a negative M_{jit}.

Assuming that spot rates are semi-annually compounded (they are calculated with the yields of government bonds), we define the discount factor by:

$$d_{jit} = \frac{1}{(1 + r_{jit}/2)^{2*T}} \tag{3}$$

where r_{jit} is the proxy for the risk-free rate associated with each particular option and T is the time span to be discounted (expressed in days). The risk-free rate associated with each option is calculated as follows.

We fit a third-degree polynomial spot curve to the daily US Treasury spot rates (see *Data* Section). In this way, we obtain a unique spot curve for each day in our sample. The daily curve is then used as a daily proxy for the risk-free rate associated with each option contract, given its particular expiration date. See Figure 3 (Appendix) for an illustration.

The future value of all announced forthcoming dividends (FVD) is determined by compounding all dividends that have been declared before today's close and that go ex-dividend between tomorrow and expiration. This means that we have to compute the FVD for each option individually. We use the following equation to do so:

$$FVD_{jit} = \sum_{m=1}^{\infty} exp(1/d_{jit}^{T=\tau_{jit}-\rho_m} - 1)D_{i,m} \tag{4}$$

where $D_{i,m}$ is the m-th dividend and ρ_m is the time until the payment of the m-th dividend (in days).

Calculating Eq.(4) for the S&P500 and transforming it (to index units) such that it can be subtracted from the price (as is necessary in Eq.(2)) would be a computational burden. Fortunately, in case of European options (i.e. SPX), the put-call parity holds exactly and allows us to apply a shortcut for calculating the time-adjusted moneyness. The Forward price (F) of a dividend-paying underlying is given by:

$$F = \exp(1/d_{jit}^{T=\tau_{jit}} - 1)S_i - FVD_{jit} \tag{5}$$

The denominator in the lag term of Eq.(2) (right side of Eq.(5)) can now be substituted by F (left side of Eq.(5)). F, in its turn, is obtained by rewriting and solving the put-call parity:

$$c - p = d^{T=\tau}(F - K) \tag{6}$$

where c is the current value of the call, p the current value of the put and all other terms have their formerly defined meaning. Note that the put-call parity relationship only holds exactly for European options, so for the equity options we still have to calculate the time-adjusted moneyness according to Eq.(2).

Different from Hentschel (2003) and Bernales and Guidolin (2014)[(3)], we use OLS to estimate the deterministic *IVS* model. To prevent misleading cross-sectional regression results, we only re-estimate the deterministic *IVS* coefficients on days with more than six option contracts[(4)]. In other words, for days with six or less option contracts the coefficients are set equal to the coefficients estimated on the preceding trading day. Table 2 shows summary statistics for the (average) estimated coefficients from Eq.(1).

[(3)] Hentschel (2003) recommends the use of Generalized Least Squares (GLS). However, Bernales and Guidolin (2014) use GLS as well, but report largely similar results when applying OLS.

[(4)] Following the general convention that one requires at least $N + 1$ observations to estimate a model with N parameters.

Table 2. Estimated coefficients of the deterministic *IVS* model

Coefficient	mean	std. dev.	skew	exc. kurt.	min.	max.	t-Test	F-test
Equity Options								
β_0	-1.25	0.28	0.14	-0.64	-2.95	0.00	-116.70 (99.78%)	
β_1	-0.04	0.05	-0.02	136.45	-8.85	3.15	-4.37 (58.24%)	
β_2	-0.01	0.07	-3.41	65.50	-4.12	2.13	-0.14 (37.94%)	
β_3	-0.01	0.04	-1.84	41.74	-3.83	1.95	-1.09 (42.49%)	
β_4	-0.01	0.05	-0.21	203.69	-7.62	11.56	-0.64 (19.64%)	
								34.46 (77.36%)
R^2	0.72	0.05	-0.98	3.81	0.49	0.83		
S&P500 Index Options								
β_0	-1.67	0.28	-0.08	-0.73	-2.28	-0.89	-345.97 (100%)	
β_1	-0.14	0.05	-0.65	0.50	-0.35	0.09	-26.62 (99.49%)	
β_2	0.02	0.03	1.38	1.82	-0.03	0.15	5.42 (75.76%)	
β_3	0.03	0.05	-0.52	0.33	-0.17	0.16	5.46 (86.92%)	
β_4	-0.03	0.03	-1.13	1.26	-0.19	0.05	-2.99 (61.09%)	
								462.01 (100%)
R^2	0.83	0.15	-1.58	2.71	0.16	0.99		

Table 2: Summary statistics of the (average) estimated deterministic *IVS* model coefficients. In parenthesis are the percentages of significant test results, using $\alpha = 5\%$. Note that all values in the *Equity Options* section are averages across all 100 underlying assets and across all days in the sample. The values in the *S&P500 Index Options* section are averages across all days in the sample.

The intercept β_0 denotes the common log-volatility in a world where Black and Scholes' (1973) assumption of constant volatility holds. In such a world, the average volatility common in option contracts across all 100 underlying assets and over the time period from 1996 to 2006 would be equal to $exp(-1.25) = 28.65\%$. For S&P500 options this value is quite a bit lower; $exp(-1.67) = 18.82\%$. Of course, it is a very explicable observation that S&P500 (implied) volatility is significantly lower than the average (implied) volatility of 100 of its constituents; after all, the S&P500 is a weighted collection of those 100 (and 400 other) assets. In this way, the S&P500 diversifies the risks (i.e. volatility) associated with the individual constituents—resulting in a lower (implied) volatility. β_1 and β_3 describe the moneyness and maturity slope respectively. β_2 defines the smile curvature in the moneyness dimension. Any moneyness and time-to-maturity interplay is captured by β_4. As Table 2 shows, the parameters of Eq.(1) are extremely often significant in case of index options. Regarding equity options, the significance percentages are lower; although β_0 is significant at nearly all times, the β_2, β_3 and β_4 coefficients are significant in only less than half of the time. An indication for in-sample fit of Eq./model (1) is given by R^2. The range of minimum and maximum values for R^2 (49%-83% for equities and 16%-99% in case of the index) imply that there is a vast variation in the explanatory power of our model. The F-test indicates a well-specified model in 77.36% of all daily regressions for all equity options (not averaged), and in 100% of all daily regressions regarding index options. Figure 2 shows a time series plot of the (average) daily deterministic *IVS* parameters.

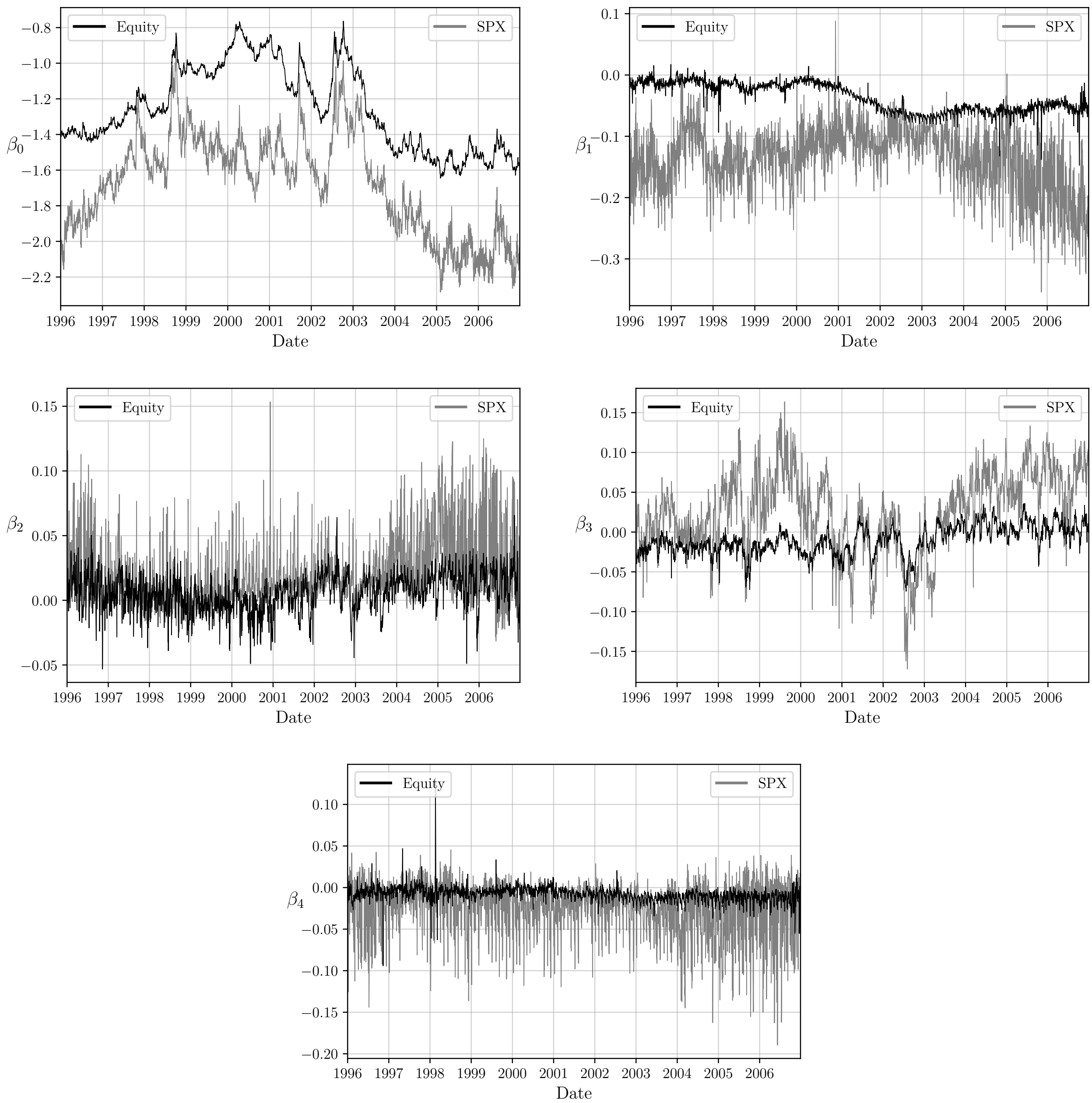

Figure 2: Time series plot of (in gray) the estimated daily coefficients of Eq.(1) for SPX and (in black) the average estimated daily coefficients for equity options (across all 100 underlyings). Data spans the time period from January 4, 1996 to December 29, 2006.

Figure 2 shows a clear correlation between the average intercept coefficient (β_0) time series for equity options and the intercept coefficient time series for index options. Correlation between β_1, β_2 and β_4 coefficients seems less evident. β_3 clearly moves in a more synchronous way (particularly noticeable during 1999-2000 and 2002-2003 time period), although movement in the average equity β_3 time series is often smaller of magnitude. In order to further investigate the joint dynamics of equity and SPX textitIVS parameters, we run 100 separate correlation analyses between the time series *IVS* coefficients (Eq.(1)) of the index and each set of equity options for an individual underlying asset. More knowledge of this correlation structure is desirable in order to propose appropriate models. Table 3 shows that, statistically, we observe a significant relationship in 69%, 89% and 57% of all trading days, for β_1, β_2 and β_4 respectively.

Table 3. Correlation analysis

	β_0^{Eq}	β_1^{Eq}	β_2^{Eq}	β_3^{Eq}	β_4^{Eq}	β_0^{SPX}	β_1^{SPX}	β_2^{SPX}	β_3^{SPX}	β_4^{SPX}
β_0^{Eq}	1 (100%)									
β_1^{Eq}	0.189 (93%)	1 (100%)								
β_2^{Eq}	-0.287 (100%)	-0.047 (87%)	1 (100%)							
β_3^{Eq}	-0.447 (99%)	-0.026 (82%)	0.451 (100%)	1 (100%)						
β_4^{Eq}	0.032 (73%)	0.116 (86%)	0.193 (96%)	0.127 (80%)	1 (100%)					
β_0^{SPX}	0.719 (100%)	0.347 (87%)	-0.286 (74%)	-0.379 (100%)	0.102 (71%)	1 (100%)				
β_1^{SPX}	0.139 (97%)	0.072 (69%)	-0.118 (57%)	-0.005 (91%)	0.075 (54%)	0.386 (100%)	1 (100%)			
β_2^{SPX}	-0.079 (99%)	-0.044 (87%)	0.095 (89%)	-0.013 (87%)	-0.063 (67%)	-0.436 (100%)	0.322 (100%)	1 (100%)		
β_3^{SPX}	-0.277 (98%)	-0.138 (63%)	0.084 (52%)	0.277 (98%)	-0.027 (47%)	-0.595 (100%)	-0.443 (100%)	0.055 (100%)	1 (100%)	
β_4^{SPX}	0.059 (96%)	0.042 (75%)	-0.058 (75%)	-0.028 (40%)	0.047 (57%)	0.194 (100%)	-0.409 (100%)	-0.860 (100%)	-0.002 (100%)	1 (100%)

Table 3: The table shows average values of separate correlation analyses between the time series *IVS* coefficients (Eq.(1)) of index options (one time series) and each individual set of equity options on a particular underlying asset (100 time series). So, we performed 100 individual correlation analyses to produce the results presented here. The parentheses show the percentage of significant correlations at a 5% significance level.

Correlations between all deterministic *IVS* coefficients are presented in Table 3. On average, many significant correlations between *IVS* parameters of equity and index options exist: for example, $\beta_0^{Eq} - \beta_0^{SPX}$ and $\beta_3^{Eq} - \beta_0^{SPX}$ are significantly serially correlated for every equity (100%). Another five coefficient-combinations report significant correlation for more than 90 out of 100 equities. Additionally, correlations between equity coefficients and their lags (known as *autocorrelations*) are statistically strong, as reported in Figure 4.

3.2 Joint cross-sectional dynamics of equity-SPX *IVS*

The focus of the research is on the joint dynamics of individual equity *IV* surfaces and the SPX *IVS*. Bernales and Guidolin (2014) report significant evidence for (serial) correlation on the same data that is used in this research. We find similar proof based on the results in Figure 2, Table 3 and Figure 4 (Appendix). All together, this legitimizes the choice for a VAR-type model. Table 3 also justifies the additional inclusion of exogenous lags, due to the strong correlations. Based on the preceding findings, we follow the paper of Bernales and Guidolin (2014) and apply a dynamic equity-SPX *IVS* model (abbreviation: *VARX-SPX*):

$$\hat{\beta}_t^{Eq} = \gamma + \sum_{j=1}^{p} \Phi_j \hat{\beta}_{t-j}^{Eq} + \sum_{k=1}^{q} \Psi_k \hat{\beta}_{t-k}^{SPX} + u_t \tag{7}$$

where $u_t \overset{iid}{\sim} N(0, \Omega)$ and $\hat{\beta}_t^{Eq} = [\hat{\beta}_{0,t}^{Eq}, \hat{\beta}_{1,t}^{Eq}, \hat{\beta}_{2,t}^{Eq}, \hat{\beta}_{3,t}^{Eq}, \hat{\beta}_{4,t}^{Eq}]$ is the time series of estimated parameters from the deterministic *IVS* model at t for underlying equity (Eq). $\hat{\beta}_t^{SPX}$ represents the time series of estimated parameters from the deterministic *IVS* model at t for the index options. The intercept γ is of size 5×1 and Φ_j and Ψ_k are of size 5×5. In this way, the maximum number of potential parameters is 155(5). The number of lags p and q are determined by minimization of the Bayes-Schwarz criterion (BIC) on the integer interval $p, q \in [1, 3]$.

(5) $[1 \times (5 \times 1) + p \times (5 \times 5) + q \times (5 \times 5)]$ with $p, q \in [1, 3]$

3.3 Benchmark models

For comparative reasons, we also evaluate three other models. The first of these benchmark models is partly equal to the VARX(p,q) model, except that we force $q = 0$, resulting in:

$$\hat{\beta}_t^{Eq} = \delta + \sum_{j=1}^{p} \Theta_j \hat{\beta}_{t-j}^{Eq} + \upsilon_t \tag{8}$$

where $\upsilon_t \overset{iid}{\sim} N(0, \Xi)$ and $\hat{\beta}_t^{Eq} = [\hat{\beta}_{0,t}^{Eq}, \hat{\beta}_{1,t}^{Eq}, \hat{\beta}_{2,t}^{Eq}, \hat{\beta}_{3,t}^{Eq}, \hat{\beta}_{4,t}^{Eq}]$ is the time series of estimated parameters from the deterministic *IVS* model at t for underlying equity (Eq). This model does not consider past information on the S&P500 index IVS, and is essentially just a simple VAR(p) model. The intercept δ is of size 5×1 and Θ_j is of size 5×5. In this way, the maximum number of potential parameters is 80[(6)]. Again, the number of lags p are determined by minimization of the Bayes-Schwarz criterion (BIC) on the integer interval $p \in [1, 3]$.

Secondly, we evaluate an ad-hoc Strawman model. It has been tested by Dumas et al. (1998) and used as a benchmark by Christoffersen and Jacobs (2004) and Goncalves and Guidolin (2006, 2014). The model fits a random walk to the *IVS* coefficients in the deterministic model. The best prediction for tomorrow is given by the random walk law of motion:

$$\hat{\beta}_t^{Eq} = \hat{\beta}_{t-1}^{Eq} \tag{9}$$

The third benchmark model also applies a random walk, but instead on implied volatilities directly. It is described by Harvey and Whaley (1992) and has been applied by, among others, Konstantinidi et al. (2008). Tomorrow's best prediction is given by:

$$\hat{\sigma}_t = \sigma_{t-1} \tag{10}$$

We will refer to this model as the Random Walk *IV* model.

3.4 Volatility Indexes

We extend the existing literature by proposing an alternative method of incorporating information from the market *IVS* in the volatility surface of equity options. Instead of adding estimated deterministic SPX *IVS* parameters (Eq.(1)) to create a dynamic equity-SPX *IVS* model (Eq.(7)), we add a time series of the *Volatility Index Term Structure* (VIX Term Structure) data. VIX Term Structure is the term used by the Chicago Board Options Exchange (CBOE) for a set of implied/expected volatilities in the S&P500 over different maturities. We create and use a VIX term structure consisting of daily close prices of four different *Volatility Indexes* (VI); the VXST, VIX[(7)], VXV and VXMT. These volatility indexes are estimates of implied/expected volatility in the S&P500 over a 9-day, 30-day/1-month, 3-month and 6-month time horizon, respectively, and are calculated every 15 seconds throughout the trading day. The values for the indexes are computed by the CBOE using the midpoint of real-time SPX bid/ask quotes for options within a predetermined range of strikes and maturities (for more information about the computation aspect, we refer to the Volatility Indexes documentation on the official CBOE website). The historical data on the Volatility Indexes is downloaded from Yahoo Finance.

(6) $[1 \times (5 \times 1) + p \times (5 \times 5)]$ with $p \in [1, 3]$

(7) Note that the VIX index is only one value measuring implied volatility over 1-month ahead, whilst VIX Term Structure is a set of several implied volatility values over different time horizons, including the 1-month time horizon. VIX should not be confused with the VIX Term Structure

Contrary to the dynamic equity-SPX *IVS* model of Goncalves and Guidolin (2006) and Bernales and Guidolin (2014), this proposed model does not require the estimation of a deterministic *IVS* model for SPX. Instead, we construct a dynamic equity *IVS* - Volatility Indexes model (abbreviation; *VARX(p,q)-VI*):

$$\hat{\beta}_t^{Eq} = \varphi + \sum_{j=1}^{p} \Delta_j \hat{\beta}_{t-j}^{Eq} + \sum_{k=1}^{r} \Upsilon_k VI_{t-k} + \nu_t \tag{11}$$

where $\nu_t \overset{iid}{\sim} N(0, \Sigma)$, $\hat{\beta}_t^{Eq} = [\hat{\beta}_{0,t}^{Eq}, \hat{\beta}_{1,t}^{Eq}, \hat{\beta}_{2,t}^{Eq}, \hat{\beta}_{3,t}^{Eq}, \hat{\beta}_{4,t}^{Eq}]$ is the 5×1 vector time series of estimated parameters from the deterministic *IVS* model at t for some underlying equity (Eq) and $VI_t = [VXST_t, VIX_t, VXV_t, VXMT_t]$ is the 4×1 vector time series of daily close prices for CBOE's SPX related volatility indexes (together forming a VIX Term Structure) on day t. Because β_t^{Eq} are obtained from a log implied volatility model, we also take the log of the VIX Term Structure variables before we include them in the model (Eq.(eq:extensionVARX)). The intercept φ is of size 5×1, Δ_j is of size 5×5 and Υ_j is of size 5×4. In this way, the maximum number of potential parameters is 140[(8)]. Again, the number of lags (p, q) are determined by minimization of the Bayes-Schwarz criterion (BIC) on the integer interval $p, q \in [1, 3]$.

To the best of our knowledge, there is no existing, published literature on this topic. The main reason for this might be the fact that most of these indexes are relatively new and therefore the sets of historical observations have been scarce. The VIX dates from 1990, but the VXV, VXMT and VXST only date from 2007, 2008 and 2011, respectively. Due to this fact, we back-test our proposed extension model on the post-financial crisis time period from January 3, 2011 to December 31, 2015 (1258 trading days). Although the focus of this paper is the time period from January 4, 1996 to December 29, 2006, in order to compare the extension model fairly, the other models in this paper have also additionally been estimated and forecasted on the latter time period (2011-2015). For this purpose, all collected data as described in the *Data* section, is also downloaded for the 2011-2015 time period. The selection of stocks for the 2011-2015 time period comprises of 90 (see **bold** tickers in Appendix B), as for some of the 100 originally selected companies we were unable to retrieve information from WRDS.

(8) $[1 \times (5 \times 1) + p \times (5 \times 5) + q \times (5 \times 4)]$ with $p, q \in [1, 3]$

4 Results

We have two separate estimation samples:

- January 4, 1996 to December 29, 2006 (2768 trading days). Estimated models: VARX-SPX, VAR, ad-hoc Strawman, Random Walk *IV*
- January 3, 2011 to December 31, 2015 (1258). Estimated models: VARX-VI, VARX-SPX, VAR, ad-hoc Strawman, Random Walk *IV*

4.1 Implied Volatility

First, we recursively estimate the parameters of the VARX(p,q)-VI model ($\hat{\varphi}$, $\hat{\Delta}$ and $\hat{\Upsilon}$), dynamic VARX(p,q)-SPX model ($\hat{\gamma}$, $\hat{\Phi}$ and $\hat{\Psi}$) and VAR(p) model ($\hat{\delta}$ and $\hat{\Theta}$) at daily frequency, by using a rolling window of 126 days (assuming there are 252 trading days in a year). The forecast section therefore covers the following time periods:

- July 8, 1996 to December 29, 2006 (2640 trading days)
- July 6, 2011 to December 31, 2015 (1131 trading days)

By plugging in the relevant lags (p, q) of the deterministic *IVS* coefficients in Eq.(11), (7), (8) and (9), we obtain one-day-ahead forecasts for the deterministic *IVS* parameters in Eq.(1) of the VARX(p,q)-VI, VARX(p,q)-SPX, VAR(p) and Strawman model, respectively.

Now that we have several one-day-ahead forecasts for the deterministic *IVS* model using (11) the VARX(p,q)-VI model, (7) dynamic VARX(p,q) model, (8) VAR(p) model and (9) Strawman model, we plug them back into the deterministic *IVS* model (Eq.(1)). Then, for each option in our sample today, the one-day-ahead forecast of implied volatility for the identical option tomorrow is then given by inputting todays time-adjusted moneyness and time-to-maturity in the deterministic *IVS* model and taking the exponent (since it models log implied volatility):

$$\hat{\sigma}^{Eq}_{t+1} = exp\big(\hat{\beta}^{Eq}_{0,t+1} + \hat{\beta}^{Eq}_{1,t+1}M_t + \hat{\beta}^{Eq}_{2,t+1}M_t^2 + \hat{\beta}^{Eq}_{3,t+1}\tau_t + \hat{\beta}^{Eq}_{4,t+1}(M_t\tau_t)\big) \tag{12}$$

The benchmark model described by Harvey and Whaley (1992), is extremely easy in implementation. The predicted implied volatility of an option is simply given by its implied volatility today: $\hat{\sigma}^{Eq}_{t+1} = \sigma_t$.

Ultimately, we have 5 (not by definition different) predictions for the implied volatility tomorrow of each option today.

4.2 Option Prices

Using the forecasted implied volatilities obtained from the VARX and benchmark models, we can calculate the corresponding forecasted option prices. For this purpose, we use the binomial tree method of Cox et al. (1979) for American option valuation. Inputs in this model are the number of binomial steps (which we set equal to 15), underlying asset price, strike price, risk-free rate, volatility (forecasted) and time-to-maturity. Since we do not have forecasts for the underlying closing price and (proxy for the) risk-free rate, we follow Bernales and Guidolin (2014) and take the values at t. For each tradeable option today, we obtain 5 different one-day-ahead forecasts for the mid price.

4.3 Statistical measures of forecasting performance

Out-of-sample forecasting performance for all models (Eq.(11), (7), (8), (9) and (10)) is assessed using a number of conventional criteria, namely: the root mean squared prediction error (RMSE), mean absolute prediction error (MAE) and the mean correct prediction of direction of change (MCPDC) statistic. The implied volatility prediction error is the difference between forecasted implied volatility and true Black-Scholes implied volatility. In case of option prices, prediction errors are defined to be the difference between forecasted (closing) mid price and the middle of the true best closing bid and ask price. MCPDC is the average frequency for which the directional change of the forecast relative to the value at t is of the same sign as the true observed directional change. MCPDC is particularly interesting, since it is directly related to trading strategy profitability. Also, it gives an idea of the performance compared to a pure guessing strategy (in which case a MCPDC of around 50% would be expected). Table 4 shows the average performance measures in terms of implied volatility (-IV).

Table 4. Out-of-sample measures of predictability in terms of implied volatilities

	Implied Volatilities					
	January 4, 1996 - December 29, 2006			January 3, 2011 - December 31, 2015		
	RMSE-IV	MAE-IV	MCPDC-IV	RMSE-IV	MAE-IV	MCPDC-IV
VARX(p,q)-VI	NA	NA	NA	0.0253	0.0137	56.00%
VARX(p,q)-SPX	0.0322	0.0139	57.61%	0.0254	0.0138	55.65%
VAR(p)	0.0304	0.0133	57.76%	0.0250	0.0134	55.76%
Strawman	0.0251	0.0122	57.80%	0.0235	0.0130	55.42%
Random Walk IV	0.0188	0.0093	NA	0.0220	0.0104	NA

Table 4: Average out-of-sample statistical measures of predictability in the implied volatility dimension for the dynamic equity IVS - VI model (Eq.(11)), dynamic equity-SPX IVS model (Eq.(7)), VAR benchmark model (Eq.(8)), ad-hoc Strawman benchmark model (Eq.(9)) and Random Walk IV (Harvey and Whaley, 1992) benchmark model (Eq.(10)). MCPDC for the Random Walk IV model can not be calculated, since the model predicts a no-change implied volatility. Values for dynamic equity IVS - VI model can not be calculated for the period January 4, 1996 - December 29, 2006 because not all required data is available during this period.

The out-of-sample RMSE-IV for the VARX-SPX model are the worst, with values of 3.22% (1996-2006 time period) and 2.54% (2011-2015 time period). Closely following are the VARX-VI model (2011-2015: 2.53%), VAR model (1996-2006: 3.04%, 2011-2015: 2.50%) and ad-hoc Strawman model (1996-2006: 2.51%, 2011-2015: 2.35%). Interestingly, the random walk IV model described by Harvey and Whaley (1992), predicting a no-change in implied volatility, outperforms all other models, with considerably lower RMSE of just 1.88% (1996-2006) and 2.20% (2011-2015). Based on MAE and MCPDC, we observe a largely similar ranking. We think this is rather surprising, actually. Since implied volatilities possess high persistence (observable in the autocorrelation plot in Figure 4), we would have expected a better performance of more sophisticated models (such as VARX-VI, VARX-SPX or VAR) that are better able to capture this persistence than a static random walk model. Anyway, it is interesting to see that the MCPDC of all models exceeds 50%—whereas in an unpredictable environment, results close to 50% would be expected. So although the equity IVS - VI model and dynamic equity-SPX IVS model obviously do not perform the best in our back-test, they still seem capable of outperforming a solely guessing market participant (with an expected MCPDC of around 50%). However, the equivalent conclusion holds for the more parsimonious VAR and Strawman model.

Exactly the same conclusions can be drawn from the performance measures in terms of option price (-P) perspective (see Table 5). The out-of-sample RMSE for the VARX-SPX model is the worst, with a value of $0.8793 (1996-2006 time period) and $0.4524 (2011-2015 time period). Closely following are the VARX-VI model (2011-2015: $0.4512),

Table 5. Out-of-sample measures of predictability in terms of option prices

	Option Prices					
	January 4, 1996 - December 29, 2006			January 3, 2011 - December 31, 2015		
	RMSE-P	MAE-P	MCPDC-P	RMSE-P	MAE-P	MCPDC-P
VARX(p,q)-VI	NA	NA	NA	0.4512	0.2953	74.46%
VARX(p,q)-SPX	0.8793	0.2244	78.94%	0.4524	0.2961	74.44%
VAR(p)	0.8759	0.2212	79.12%	0.4508	0.2947	74.46%
Strawman	0.8691	0.2162	79.45%	0.4500	0.2935	74.73%
Random Walk IV	0.8570	0.2121	NA	0.4215	0.2707	NA

Table 5: Average out-of-sample statistical measures of predictability in the option price dimension for the dynamic equity IVS - VI model (Eq.(11)), dynamic equity-SPX IVS model (Eq.(7)), VAR benchmark model (Eq.(8)), ad-hoc Strawman benchmark model (Eq.(9)) and Random Walk IV (Harvey and Whaley, 1992) benchmark model (Eq.(10)). MCPDC for the Random Walk IV model can not be calculated, since the model predicts a no-change in implied volatility. Values for dynamic equity-VI IVS model can not be calculated for the period January 4, 1996 - December 29, 2006 because not all required data is available during this period.

VAR model (1996-2006: \$0.8759, 2011-2015: \$0.4508) and ad-hoc Strawman model (1996-2006: \$0.8691, 2011-2015: \$0.4500). Again, the most parsimonious model, predicting a no-change in implied volatilities (and thus in option prices), outperforms all other models. This holds in terms of MAE as well. With regards to the MCPDC, Strawman outperforms a VAR model and the VAR and VARX-VI models outperform the VARX-SPX model. Again, note that all MCPDC-P values are far greater than 50%, indicating directional predictability to at least some extent. Under these circumstances, one might exploit the directional change predictions to generate profits from purely directional triggered trading strategies. But even simple models appear to be sufficient for this purpose.

One particular interesting observation from Table 5 is the fact that RMSE's of the 2011-2015 sample are nearly half of the RMSE's of the 1996-2006 sample, while MAE and MCPDC remain more or less in the same order of magnitude. The reason for this probably is related to the difference in stocks selection (1996-2006 all stocks in Appendix B, 2011-2015 only bold tickers in Appendix B) and different time period. Another striking difference, between Table 4 and 5, is the height of MCPDC values. In terms of option prices average MCPDC is 56.57%, whilst in terms of implied volatilities average MCPDC equals 76.51%. This means that there is a percentage of times when our directional estimate in terms of implied volatility is opposite from the true, observed directional change in implied volatility, but at the same time our directional option price forecast (directly following from our volatility forecast and the Black-Scholes formula) is equal to the true, observed directional price change. Given the difference in MCPDC between implied volatilities and option prices, it must be that the direction of our option price forecasts is more often equal to the true directional change than this is the case regarding our implied volatility forecasts. However, everything else equal, the Black-Scholes formula assigns higher (lower) prices when implied volatilities are higher (lower). This implies that the direction of our forecasts is always equal to the direction of our option price forecasts. Henceforth, the fact that we are more often correct regarding price forecasts can only be attributed to the fact that the other inputs in the Black-Scholes model (such as underlying price at $t + 1$) more often than not have shifted the directional change in option price in favor of our forecast. Most of the times this will probably happen when forecasted volatilities for $t + 1$ are close to observed volatilities at t.

We also want to determine which model produces the most accurate forecasts. The Diebold-Mariano (1995) test is ideal for this purpose. The test has a null-hypothesis of equal forecast accuracy given some loss function. There is no reason to assume that symmetry between positive and negative forecast errors is inappropriate, so we use a symmetric loss function defined as the differences between squared prediction errors.

Table 6. Equal predictive accuracy (Diebold-Mariano) test of **NA** model against the others:

	January 4, 1996 - December 29, 2006				January 3, 2011 - December 31, 2015			
	Implied Volatilities		Option Prices		Implied Volatilities		Option Prices	
	value	sign.	value	sign.	value	sign.	value	sign.
VARX(p,q)-VI	NA	NA	NA	NA	**NA**	**NA**	**NA**	**NA**
VARX(p,q)-SPX	**NA**	**NA**	**NA**	**NA**	-1.42	49%	-2.82	64%
VAR(p)	3.44	72%	5.66	83%	4.24	67%	1.12	44%
Strawman	6.82	72%	7.47	83%	6.13	67%	2.31	44%
Random Walk *IV*	10.99	89%	9.71	87%	7.76	93%	28.98	100%

Table 6: Average values for the Diebold-Mariano test statistic across all individual underlying assets and over time and the percentage of statistically significant test results (at 5% significance level, where DM $\sim N(0,1)$). In case of a significant result, the null hypothesis of equal forecast accuracy is rejected. **For the period January 4, 1996 - December 29, 2006, we compare all models to the VARX-SPX model**. It can not be compared to itself, so in that case table entries are not available (**NA**). Additionally, it can not be compared to the VARX-VI model, because it is not possible to estimate the VARX-VI model during this time period due to a lack of information (NA). **For the period January 3, 2011 - December 31, 2015, we compare all models to the VARX-VI model**. It can not be compared to itself, so in that case table entries are not available (**NA**).

For positive DM-values the null hypothesis is rejected in favor of the corresponding benchmark model (1996-2006: VAR, Strawman and Random Walk *IV*, 2011-2015: VARX-SPX, VAR, Strawman and Random Walk *IV*). In case of negative values, the hypothesis is rejected in favor of the dynamic equity-SPX *IVS* (1996-2006) or VARX-VI model (2011-2015). Both in terms of implied volatility and option price forecasts, the DM tests most frequently rejects the null hypothesis of equal predictive accuracy in favor of the more parsimonious models. For the time period 1996-2006 it is clear that the VARX-SPX model has a lower forecast accuracy than the VAR, Strawman and Random Walk *IV* model. Regarding the 2011-2015 time period, it becomes clear that the hypothesis of equal forecast accuracy between the VARX-VI and VAR, Strawman and Random Walk *IV* model is rejected in an overwhelmingly percentage of times (89%, 87%, 93% and 100%). However, average values for the DM statistic are more conservative for the VARX-VI versus VARX-SPX model (-1.42 and -2.82 in implied volatility and option price dimension, respectively). Also the significance percentages (49% and 64%) show that the VARX-VI and VARX-SPX models perform rather similar in terms of forecast accuracy.

Based on all above findings, we do not find significant evidence that, on average, a VARX-SPX or VARX-VI model is superior to any of the benchmark models (VAR, Strawman, Random Walk *IV*. Instead, it is proved that the more parsimonious a model is the lower its values for RMSE and MAE and the higher its value for MCPDC are. The MCPDC-P values indicate directional predictability to at least some extent. However, this is not necessarily reserved for more complex models; even the more parsimonious models (i.e. VAR, Strawman) are sufficient to generate forecasts that well exceed a 50% correct directional change frequency. Therefore, it makes no sense and is not worth the additional effort to use the more complex models. An explanation for the relatively well performance of the Random Walk *IV* and Strawman models would be that the data generating process underlying the implied volatilities has a tendency to follow a randomized path (on an inter-day basis). This supposition is supported by the relatively weak performance of models incorporating vector (exogenous) autoregressive terms (VARX-VI, VARX-SPX, VAR). Further research is required to prove whether this also holds for intra-day/higher frequency data.

It should be stated that our findings are at odds with the results of Bernales and Guidolin (2014). Although the approach in our paper is largely similar, we introduced the following simplifications. (1) Instead of 150 stocks, we included 100 stocks. However, the inclusion of 50 randomly excluded stocks is unlikely to alter the average model performances as presented in this paper drastically. (2) We used GLS instead of OLS. Nevertheless, Bernales and Guidolin (2014) themselves already note only a small difference between OLS and GLS results. And (3) we did not account for dividends in the no-arbitrage bounds calculation. But although Bernales and Guidolin (2014) do not specify the definition of the bounds they apply, a potential small alteration in specification, i.e. inclusion of dividends, would certainly not significantly influence the main results in this paper, as dividends are sparse considering datasets spanning several years in trading days. Another difference might be in the calculation of the discount rate, yet we can not know for sure because Bernales and Guidolin (2014) do not mention how they calculate it. Probably the biggest potential cause of the contradicting results is in the handling of days with too few observations (see *Methodology* Section); but again the approach of Bernales and Guidolin regarding this subject is not clearly stated. Considering all the above, the true cause of the discrepancy in main results remains largely unclear.

5 Conclusion and discussion

This paper has focused on forecasting the dynamics of equity IV surfaces at daily frequency. Bernales and Guidolin (2014) suggest that the difference in trading frequency may indicate that changes in IVS are more rapidly integrated into prices for SPX than for equity options—and thus, they reason, SPX IVS dynamics may be exploited to predict equity IVS dynamics. In order to investigate these joint dynamics, a VARX model has been adopted. However, despite the suggestions and concluding results of Bernales and Guidolin (2014), this paper has found contradicting evidence—models incorporating exogenous autoregressive dynamics are outperformed in terms of prediction error and forecast accuracy measures by more parsimonious, static models (Strawman, Random Walk IV). Nonetheless, the presence of many significant correlations between equity and SPX IVS parameters is proven, indicating that the IV surfaces of equity and SPX options are in fact dynamically related to some extent. Yet, correlation does not imply causality—demonstrated by the fact that this research fails to prove significant added value in modeling equity IVS dynamics by incorporating (historical) information on SPX IVS dynamics. Besides the empirical evidence, we also question the reasoning behind the motive of Bernales and Guidolin (2014). Even with the existence of discrepancies in trading frequencies, this should not at all necessarily lead to out-dated IV surfaces for equity options; also in the absence of trading volume, bid/ask quotes may be continuously updated and might be priced 'correctly' (no information bias with regards to SPX option quotes). Additionally, this paper studied the explanatory value of CBOE's volatility indexes for equity IVS dynamics. Again, the proposed model did not provide significant evidence of out-performing more parsimonious, static models (Strawman, Random Walk IV). However, its performance has proven fairly equal to (at least not worse than) the dynamic equity-SPX IVS model of Bernales and Guidolin (2014), but does not require the estimation of the deterministic IVS model for SPX. We would not rule out the existence of information asymmetry in equity and SPX option quotes due to liquidity biases completely, however. But at least, such a bias may more likely be expected to be present in higher frequency data (intra-day instead of inter-day). It would be interesting for further research to focus on higher frequency dynamics.

References

[1] Bakshi, G., Cao, C., and Chen, Z. (2000). "Do call prices and the underlying stock always move in the same direction?", *Review of Financial Studies*, 13 (3): 549-84.

[2] Bernales, A., and Guidolin, M. (2014). "Can we forecast the Implied Volatility Surface Dynamics of equity options? predictability and economic value tests", *Journal of Banking & Finance*, 46:326-342.

[3] Black, F. (1976). "Studies of stock price volatility changes", *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistic Section*, 177-181.

[4] Bloch, D. (2012). "From Implied Volatility Surface to Quantitative Options Relative Value Trading" (September 7, 2012). Available at SSRN: https://ssrn.com/abstract=2143101 or http://dx.doi.org/10.2139/ssrn.2143101

[5] Borovkova, S.A., and Permana, F.J. (2009). "Implied volatility in oil markets", *Computational Statistics and Data Analysis*, vol 53, no. 6, pp. 2022-2039. DOI: 10.1016/j.csda.2008.02.013

[6] Burke, D. (1988). "Estimation of a bivariate distribution function under random censorship", *Biometrika*, 75 (2): 379-82.

[7] Campa, J., and Chang, K. (1995). "Testing the Expectations Hypothesis on the Term Structure of Volatilities", *Journal of Finance*, 50, 529-547.

[8] Christoffersen, P., and Jacobs, C. (2004). "The importance of the loss function in option valuation", *Journal of Financial Economics*, 72, 291-318.

[9] Christoffersen, P., Fournier, M., and Jacobs, K. (2016). "The factor structure in equity options" (September 13, 2016), *Rotman School of Management* Working Paper No. 2224270.

[10] Black, F., and Scholes, M. (1973). "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 637-654.

[11] Cox, J., Ross, S., and Rubinstein, M. (1979). "Option Pricing: A Simplified Approach", *Journal of Financial Economics*, 7, 229—263.

[12] Diebold, F.X., and Mariano, R.S. (1995). "Comparing Predictive Accuracy". *Journal of Business and Economic Statistics*, 13: 253-63.

[13] Dumas, B., Fleming J., and Whaley R. (1998). "Implied Volatility Functions: Empirical Tests", *Journal of Finance*, 53, 2059-2106.

[14] Figlewski, S., (1989). "Options Arbitrage in Imperfect Markets", *Journal of Finance*, 44, 1289-1311.

[15] Goncalves, S., and Guidolin, M. (2006). "Predictable Dynamics in the S&P 500 Index Options Implied Volatility Surface", *Journal of Business*, 79 (3), 1591-1636.

[16] Harvey, C., and Whaley, R. (1992). "Market Volatility Prediction and the Efficiency of the S&P 100 Index Options Market", *Journal of Financial Economics*, 31, 43-73.

[17] Hentschel, L. (2003). "Errors in implied volatility estimation", *Journal of Financial and Quantitative Analysis*, 38, 779–810.

[18] Heston, S. (1993). "A closed-form solution for options with stochastic volatility with applications to bond and currency options", *Review of Financial Studies*, 6 (2): 327-43.

[19] Heston, S., and Nandi, S. (2000). "A Closed-Form GARCH Option Valuation Model", *Review of Financial Studies*, 13, 585-625.

[20] Konstantinidi, E., Skiadopoulos, G., and Tzagkaraki, E. (2008). "Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices", *Journal of Banking & Finance*, 32 (11): 2401-11.

[21] Rubinstein, M. (1994). "Implied Binomial Trees", *Journal of Finance*, 49, 781-818.

[22] Sun, S., and Ji, S. (2015). "Estimation of implied volatility surface and its dynamics: evidence from S&P500 index option in post-financial crisis market", Master of Science in Finance thesis, Simon Fraser University, Canada.

[23] Tompkins, R., D'Ecclesia, R.L. (2006). "Unconditional Return Disturbances: A Non Parametric Simulation Approach", *Journal of Banking and Finance*, 30 (1), 287-314.

A Appendix: figures

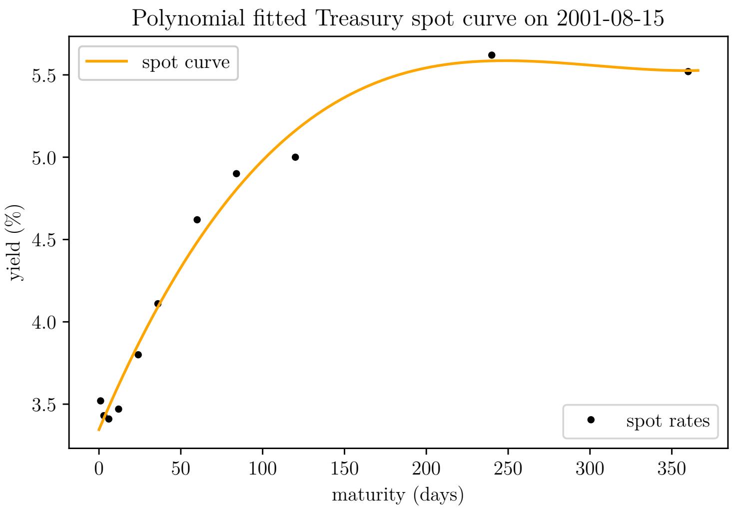

Figure 3: US Treasury spot rates and the fitted third-degree polynomial spot curve on 2001-08-15. The proxy for the risk-free rate associated with an option contract on this date is given by the y-coordinate of the curve at the days-to-maturity (corresponding to the option) x-coordinate. The spot curve is unique for each date.

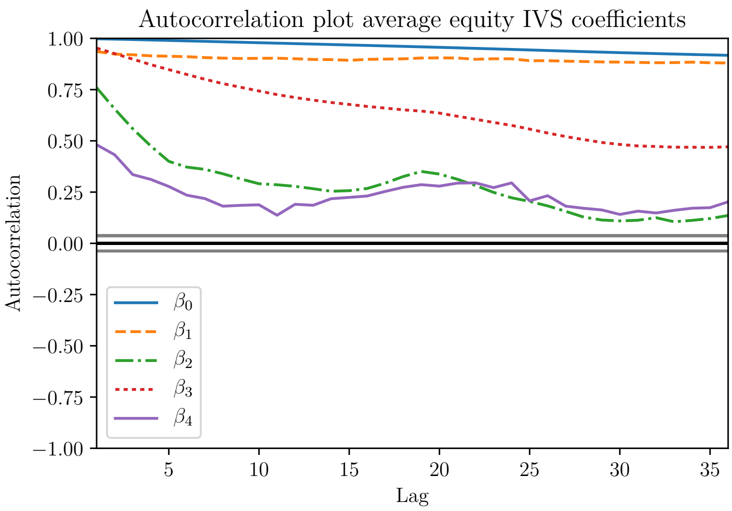

Figure 4: Correlation plot of (lags of) average deterministic equity *IVS* coefficients. Gray bounds denote the 95%-significance interval of (H0:) no-correlation (rejected outside this bound).

B Appendix: list of equities

Ticker	Company Name
MMM	3M CO
ALL	ALLTEL CORP
MO	ALTRIA GROUP INC
AXP	AMERICAN EXPRESS CO
AIG	AMERICAN INTERNATIONAL GROUP
AMGN	AMGEN INC
AAPL	APPLE INC
ASH	ASHLAND GLOBAL HOLDINGS INC
ADP	AUTOMATIC DATA PROCESSING
BHI	BAKER HUGHES INC
BLL	BALL CORP
BAC	BANK OF AMERICA CORP
BCR	BARD (C.R.) INC
BDK	BLACK & DECKER CORP
HRB	BLOCK H & R INC
BSX	BOSTON SCIENTIFIC CORP
BMY	BRISTOL-MYERS SQUIBB CO
CVX	CHEVRON CORP
CI	CIGNA CORP
C	CITIGROUP INC
KO	COCA-COLA CO
CL	COLGATE-PALMOLIVE CO
CMA	COMERICA INC
CAG	CONAGRA BRANDS INC
COP	CONOCOPHILLIPS
COST	COSTCO WHOLESALE CORP
CMI	CUMMINS INC
CVS	CVS HEALTH CORP
DE	DEERE & CO
DIS	DISNEY (WALT) CO
D	DOMINION RESOURCES INC
DJ	DOW JONES & CO INC
DD	DU PONT (E I) DE NEMOURS
ETN	EATON CORP PLC
EMR	EMERSON ELECTRIC CO
XOM	EXXON MOBIL CORP
FMCC	FEDERAL HOME LOAN MORTG CORP
FDX	FEDEX CORP
GD	GENERAL DYNAMICS CORP
GE	GENERAL ELECTRIC CO
HAL	HALLIBURTON CO
HSY	HERSHEY CO
HES	HESS CORP
HPQ	HP INC
ITW	ILLINOIS TOOL WORKS
IR	INGERSOLL-RAND PLC
INTC	INTEL CORP
IBM	INTL BUSINESS MACHINES CORP
JNJ	JOHNSON & JOHNSON
KMB	KIMBERLY-CLARK CORP

Ticker	Company Name
LB	L BRANDS INC
LLY	LILLY (ELI) & CO
LNC	LINCOLN NATIONAL CORP
LMT	LOCKHEED MARTIN CORP
L	LOEWS CORP
MRO	MARATHON OIL CORP
MCD	MCDONALD'S CORP
MDT	MEDTRONIC PLC
MRK	MERCK & CO
MSFT	MICROSOFT CORP
TAP	MOLSON COORS BREWING CO
MS	MORGAN STANLEY
NEM	NEWMONT MINING CORP
NEE	NEXTERA ENERGY INC
NKE	NIKE INC
NSC	NORFOLK SOUTHERN CORP
NOC	NORTHROP GRUMMAN CORP
PCAR	PACCAR INC
PH	PARKER-HANNIFIN CORP
PEP	PEPSICO INC
PFE	PFIZER INC
PD	PHELPS DODGE CORP
PBI	PITNEY BOWES INC
PNC	PNC FINANCIAL SVCS GROUP INC
PPG	PPG INDUSTRIES INC
PX	PRAXAIR INC
PG	PROCTER & GAMBLE CO
RTN	RAYTHEON CO
SGP	SCHERING-PLOUGH
SLB	SCHLUMBERGER LTD
SHW	SHERWIN-WILLIAMS CO
STI	SUNTRUST BANKS INC
SYY	SYSCO CORP
TGNA	TEGNA INC
TIN	TEMPLE-INLAND INC
THC	TENET HEALTHCARE CORP
TXN	TEXAS INSTRUMENTS INC
TXT	TEXTRON INC
HD	THE HOME DEPOT INC
TJX	TJX COMPANIES INC
TYC	TYCO INTERNATIONAL PLC
X	UNITED STATES STEEL CORP
UTX	UNITED TECHNOLOGIES CORP
UNH	UNITEDHEALTH GROUP INC
VZ	VERIZON COMMUNICATIONS INC
WMT	WAL-MART STORES INC
WY	WEYERHAEUSER CO
WHR	WHIRLPOOL CORP
WWY	WRIGLEY (WM) JR CO
WYE	WYETH

C Appendix: list of abbreviations

API	Application Programming Interface
BIC	Bayes-Schwarz criterion (lag selection)
CBOE	Chicago Board Options Exchange
DM	Diebold-Mariano (test statistic)
GLS	generalized least squares (regression method)
IV	implied volatility
IVS	implied volatility surface
MAE	mean absolute prediction error
MCPDC	mean correct prediction of direction of change
OLS	ordinary least squares (regression method)
RMSE	root mean squared prediction error
SPX	symbol for option contracts on the S&P500 index
VAR	vector autoregression model
VARX	exogenous vector autoregression model
VI	Volatility Index
WRDS	Wharton Research Data Services