# VOLATILITY FORECASTING WITH HIGHER MOMENTS OF THE OPTION IMPLIED VOLATILITY SKEW

Evidence from S&P 500 option markets

Lappeenranta–Lahti University of Technology LUT

Master’s thesis in Strategic Finance and Business Analytics

2022

Matias Kujala

Examiners: Associate Professor Jan Stoklasa
Visiting Researcher Tomáš Talášek

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT School of Business and Management, Business Administration

Matias Kujala

**Volatility Forecasting with Higher Moments of the Implied Volatility Skew: Evidence from the S&P 500 Option Markets**

Master's thesis, 2022

85 pages, 7 figures, 7 tables and 15 appendices

Examiners: Associate Professor Jan Stoklasa, Visiting Researcher Tomáš Talášek

Keywords: Volatility forecasting, volatility smile, risk-neutral skewness, risk-neutral kurtosis, HAR-RV

The objective of this master's thesis is to examine whether the shape of the volatility smile implied by index options can be used to improve the accuracy of volatility forecasts. This thesis contains quantitative empirical research of S&P 500 index volatility and the volatility smile of its options. Volatility is forecasted with Heterogenous Autoregressive model for Realized Volatility, HAR-RV, introduced by Corsi (2003). In addition to empirical research, extensive review of previous literature on HAR-RV, volatility forecasting and the implied volatility smile is provided.

Four research questions are formulated to consider forecasting performance over different time horizons and distinct market periods. Pecking order of risk-neutral skewness and risk-neutral kurtosis is also examined as a source of additional information regarding future volatility. Daily implied volatility smiles are constructed with quadratic volatility functions, which can be thought to represent the expected future price distribution of the underlying index. Thus, first and second derivatives of quadratic volatility functions can be used as proxies for higher moments of the expected price distribution. (Bakshi, Kapadia & Madan, 2003).

Empirical results obtained suggest that the shape of the implied volatility smile contains information of future volatility, particularly for weekly and monthly horizons. Forecasting accuracy of different models suggest that risk-neutral kurtosis improves forecasts more than risk-neutral skewness. Furthermore, improvements in accuracy appear to be greater during adverse market periods. Coefficient significance of fitted models imply that while both of the risk-neutral higher moments seem to have significant relationship with future volatility, this relationship tends to be significant more often for risk-neutral kurtosis.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT-kauppakorkeakoulu

Matias Kujala

**Volatility Forecasting with Higher Moments of the Implied Volatility Skew: Evidence from the S&P 500 Option Markets**

Kauppatieteiden pro gradu -tutkielma, 2022

85 sivua, 7 kuvaa, 7 taulukkoa ja 15 liitettä

Tarkastajat: Associate Professor Jan Stoklasa, Visiting Researcher Tomáš Talášek

Avainsanat: Volatiliteettiennuste, volatiliteetihymy, riski-neutraali vinouma, riski-neutraali huipukkuus, HAR-RV

Tämän pro gradu -työn tavoitteena on selvittää, voidaanko indeksioptioiden implisiittisen volatiliteettihymyn muotoa käyttää volatiliteettiennusteiden tarkkuuden parantamiseen. Tämä opinnäytetyö sisältää kvantitatiivisen empiirisen tutkimuksen S&P 500 -indeksin volatiliteetista ja sen optioiden volatiliteettihymystä. Volatiliteettia ennustetaan Corsin (2003) kehittämällä realisoituneen volatiliteetin Heterogeenisellä Autoregressio-mallilla, HAR-RV:llä. Empiirisen tutkimuksen lisäksi tarjotaan laaja katsaus aikaisempaan kirjallisuuteen HAR-RV:stä, volatiliteettiennusteista ja implisiittisestä volatiliteettihymystä.

Neljällä muodostetulla tutkimuskysymyksellä tarkastellaan mallien ennustekykyä eri aikahorisonteilla ja eri markkinajaksoilla. Riskineutraalin vinouden ja riskineutraalin huipukkuuden tärkeysjärjestystä tarkastellaan myös tulevaa volatiliteettia koskevan lisäinformaation lähteenä. Päivittäiset implisiittiset volatiliteettihymyt muodostetaan neliöllisillä volatiliteettifunktioilla, joiden voidaan ajatella edustavan taustalla olevan indeksin odotettua tulevaa hintajakaumaa. Siten neliöllisen volatiliteettifunktion ensimmäistä ja toista derivaattaa voidaan käyttää kuvaamaan odotetun hintajakauman vinoumaa ja huipukkuutta. (Bakshi, Kapadia & Madan, 2003).

Empiiriset tulokset viittaavat siihen, että implisiittisen volatiliteettihymyn muoto sisältää tietoa tulevasta volatiliteetista, erityisesti viikoittaisille ja kuukausittaisille ennusteille. Eri mallien ennustetarkkuus viittaa siihen, että riskineutraali huipukkuus parantaa ennusteita enemmän kuin riskineutraali vinouma. Lisäksi tarkkuuden parantuminen näyttää olevan suurempi epäsuotuisina markkinajaksoina. Sovitettujen mallien kertoimien tilastollinen merkittävyys viittaa siihen, että vaikka molemmilla riskineutraaleilla korkeammilla momenteilla näyttää olevan merkittävä suhde tulevan volatiliteetin kanssa, tämä suhde on merkittävä useammin riskineutraalin huipukkuuden kannalta.

**Table of contents**

# 1. Introduction

Variance of asset price around its mean is commonly thought to represent risk inherent in said asset, with this price fluctuation termed as volatility. As return and risk go hand-in-hand in financial theory, understanding and managing risks generated by volatility in particular remains as a matter of utmost importance. As such, volatility plays a central role in risk management through risk measures like Value-at-Risk, in asset management with modern portfolio theory (Markowitz, 1952) using volatility to construct diversified portfolio with certain risk level, as well as in option pricing with models such as Black-Scholes (1973) where volatility is a main component when determining option prices. As noted by for instance by Hull (2017), there has been great interest among academics and practitioners towards forecasting volatility due to its importance, which has led to development of sophisticated models and methods as means to an end.

Financial data, particularly returns and thus volatility, are known to exhibit various stylized features. Rather than being normally distributed, returns over time often show signs of fatter tails and excessive peak relative to normal distribution, a feature known as leptokurtosis. Empirical observations have been made regarding volatility clustering in financial data, with periods with certain volatility regime tend to be persistent; for instance, high volatility is often followed by subsequent high volatility. These ever-changing volatility regimes combined with non-smooth jumps in prices result in leptokurtic distributions. (Hull, 2017) Moreover, asymmetry in return volatility tends to be a regularity, as noted by Corsi (2003) and Di Matteo (2007) among others, as downward movements tend to be more pronounced than upward movements. Known as leverage effect, this feature was first noted by Black (1976) and thought to be an indication of rising risk levels of a company as their equity value declines, thus leading to higher volatility. Naturally, models for forecasting volatility are then required to be able to replicate these features.

Several methods have been developed over the years for modelling volatility. One family of popular models is GARCH, or Generalized Autoregressive Conditionally Heteroscedastic models, first developed by Bollerslev (1986) and Taylor (1986). As GARCH models conditional variance using squared lags of residuals and lags of conditional variance itself, it is able to capture volatility clustering and thus replicate some leptokurtosis. To account for leverage effect, asymmetric variations of GARCH have been developed, such as Exponential GARCH. (Nelson, 1991). Progress has also been made regarding stochastic volatility models, where volatility is generated as a random process (Taylor, 1994).

In addition to above stylized features of financial data, Di Matteo (2007) among others provides evidence of repetitive patterns and characteristics in stock prices over periods of different lengths, effect in physics known as scaling and multi-scaling. Hence, models with long-memory characteristics are increasingly growing in demand in volatility forecasting, as GARCH and short-memory stochastic volatility models are not able to replicate these features. Corsi (2003; 2009) has introduced a cascade-type model for realized volatility, where future volatility is modelled in linear relationship with partial components of past volatility observed over periods differing in time, particularly daily, weekly and monthly. Despite its simplicity, Heterogenous Autoregressive model for Realized Volatility, or HAR-RV, is able to replicate stylized features of financial data, and even though it lacks proper characteristics, empirical evidence shows that it tends to work as a long-memory model. Moreover, it has shown overperformance in forecasting accuracy over short-memory models and equal performance relative to cumbersome true long-memory models, such as ARFIMA. (Corsi, 2009). Economic rationale behind HAR-RV lies in heterogenous market hypothesis by Müller (1993), stating that heterogeneity among market participants implies that their actions create volatility over different time horizons and in distinctive manner.

Black-Scholes (1973) option pricing model developed by Fischer Black, Myron Scholes and later Robert Merton had an immense effect on option pricing and provided dealers, hedgers and other agents a sophisticated method for pricing options and other claims. Their ground-breaking work was later awarded with Nobel prize; Black unfortunately deceasing few years earlier. The model prices options with several different variables, of which latent volatility is the sole unobservable factor in the markets. Thus, the model has long been used to infer volatilities implied by option prices, which can be thought as market expectations of the average volatility of the underlying asset during the life of the option. As Lee (2004) notes, some markets in fact quote options with their volatility implied by Black-Scholes rather than actual option premiums.

With Black-Scholes, implied volatilities of options differing only in exercise prices, that is, they have the same underlying asset and time-to-maturity, are equal. However, as often is the case with economic theories and models, Black-Scholes is subject to several intrinsic assumptions not valid in the physical world. Arguably one if not the most essential is the lognormal distribution of asset prices assumed by the model, requiring constant volatility and smooth price changes void of jumps. It is well documented that neither of these hold in practice, with dynamics of jumps in prices modelled already by Press (1967) and Merton (1976). Jumps and non-constant volatility have been observed to increase probability of extreme events, leading to fat-tailed, leptokurtic return distributions discussed above. (Hull, 2017) This divergence of physical and assumed distribution has resulted in disparity between theory and practice. Particularly after the crash of 1987, commonly referred as Black Monday with stock markets declining by tens of percent's globally, option implied volatilities have been observed to form a smile-like pattern when plotted over different exercise prices, or strikes. In the presence of implied volatility smile, in- and out-of-the-money (ITM, OTM) options with strike prices higher or lower than the underlying price seem overpriced, while at-the-money (ATM) options with strikes practically equal to the underlying appear under-priced, at least relative to prices suggested by Black-Scholes. In general, implied volatility is higher for longer-maturity options due to the time-value of option; however, the smile tends to be more pronounced near the end of the option's life due to fear of jumps. (Dumas, Fleming & Whaley 1998; Hull, 2017)

Theoretical backing for the existence of the smile is then in the divergence of physical and assumed distribution. It can be argued that informed traders understand the dynamics of non-normal returns and thus the elevated probability of extreme events. This in turn increases option premiums of ITM and OTM options required by dealers and market makers, and hence also implied volatilities of said options, relevant particularly during the emergence of these extremities. Furthermore, due to the leverage effect and negative skewness of returns, smiles of equity and index options tend to be left-skewed.

Linkage between the shape of the implied volatility smile and the underlying physical distribution of future prices has been studied in the earlier literature, and empirical evidence of the relationship and its dynamics has been provided by at least by Backus and Foresi (1997), Bakshi, Kapadia and Madan (2003) and Backus, Foresi and Wu (2004). As noted in literature, the shape of the smile can be thought to represent market participants' expectations of the density of future prices during the option life and as such, could contain predictive information of price action and volatility ex ante, especially if agents are well-informed. Main components of the observed shape of the smile are its slope, corresponding to the skewness of the expected physical price distribution, and convexity, interpreted as the kurtosis of the distribution. Bakshi et al. (2003) and Backus and Foresi (2004) note that these features of the distribution are derivable in risk-neutral manner and hence have coined them risk-neutral skewness and risk-neutral kurtosis. Empirical evidence of the information content inherent in the smile regarding future returns or volatility has been presented by Dennis and Mayhew (2002), Doran, Peterson and Tarrant (2007), Ying, Zhang and Zao (2010), Yan (2011), Byun and Kim (2013) and Wong and Heaney (2017)

Turn of the decade has been a period of extreme price movements and volatility, and hence it is undeniably relevant to address the potential predictive information in the volatility smile implied by option prices. In this thesis this is achieved by enhancing HAR-RV model by Corsi (2009) with variables depicting slope and convexity of the smile. Due to limitations in data availability and interest in this thesis to study volatility during the high-volatility period in 2020, period under scrutiny in this thesis spans from January 2018 to September 2021, thus including the market crash of 2020 instigated by Covid-19 crisis. Moreover, volatility of S&P 500 index tracking 500 most valuable U.S. companies is analysed, as the market for its options is one of the most liquid in the world, thus ensuring that the smile depicts market expectations as reliably as possible. The ultimate question this thesis tries to provide answers to is:

*Does the implied volatility smile contain predictive information regarding future volatility of S&P 500 index for daily, weekly and monthly horizons?*

Previous literature has provided ambiguous evidence of where exactly the potential predictive power resides in the smile. That is, is it risk-neutral skewness or risk-neutral kurtosis of the expected underlying distribution that should be used to source information. For instance, Yan (2011) and Byun and Kim (2013) prefer skewness and find no incremental information in kurtosis, while results by Pena, Rubio and Serna (1999) suggest the relevance of kurtosis, as high-volatility periods tend to be associated with convex smiles. Wong and Heaney (2017) find no dominance for either of the higher moments. Thus, this thesis tries to shed light on the following:

*Does either of risk-neutral third and fourth moments dominate the other as a source of information regarding future volatility?*

In addition to studying the information content as such, it is essential to examine how the potential information behaves over time, which is achieved by forecasting volatility over daily, weekly and monthly periods. Wong and Heaney (2017) reported increased forecasting performance for monthly volatility, while results by Byun and Kim (2013) suggest improved accuracy for one-day-, one-week- and one-month-ahead volatility. Third question of interest is thus:

*For which future horizon does the information content improve forecasting accuracy the most?*

As the period of interest covers the most pronounced phase of extreme volatility and price development during recent years, it is highly interesting to examine how informationally effective the smile has been over extremely turbulent markets. Doran et al. (2007) and Yan (2011) have reported that the smile contains predictive information of future market spikes and crashes at least to some extent, with Yan (2011) even using the slope of the smile as a proxy for jump risk. Thus, to assess the credibility of claims by Doran et al. (2007) and Yan (2011), the final question this thesis addresses is:

*Has the information content in the smile improved forecasting accuracy over a period of extreme volatility during Covid-19 crisis?*

Remainder of this thesis is structured as follows. Theoretical framework of the thesis is presented in the next chapter, starting with overview of volatility, including both realized and implied volatility as well as the option implied volatility smile. In addition, most popular proxies for latent volatility are presented. Then, Heterogenous Autoregressive model for Realized Volatility (Corsi, 2003; Corsi, 2009) is presented, after which theoretical framework is concluded with loss functions used to measure forecasting accuracy of the models. Previous literature is reviewed in-depth regarding information content of implied volatility, HAR-RV, risk-neutral higher moments as well as their utilization in volatility forecasting, followed by an introduction to data and methodology used in the thesis. Before concluding the thesis, empirical results for daily, weekly and monthly volatility forecasts are presented, analysed and summarised.

# 2. Theoretical Framework

This chapter provides theoretical overview of both realized and implied volatility as well as the volatility smile. Deterministic volatility function framework used to represent the smile is also presented, followed by a review of most common proxies for latent volatility. After volatility itself has been discussed, HAR-RV model by Corsi (2009) is presented in-depth, together with an overview of heterogenous market hypothesis by Müller (1993). Before moving on to previous literature, loss functions RMSE and MAE used in performance evaluation are introduced.

## 2.1. Volatility

This chapter provides an overview of volatility, both historical and volatility implied by option markets, as well as the well-documented implied volatility smile. To gain further understanding of the volatility smile, Black & Scholes-option pricing model (B&S) is introduced, as one of the reasons for the existence of the smile is the violation of assumptions made in B&S. Volatility and its forecasting has long been a subject of keen interest due to its central nature; forecasting volatility is crucial for option pricing as well as risk-management, for instance risk measures such as Value-at-Risk.

Often thought as a range of all possible outcomes of an uncertain variable, volatility is commonly used as a measure of risk in financial markets, and indicative ratios regarding risk-neutral returns, such as Sharpe Ratio, often utilize realized volatility as a denominator of the ratio. However, as risk has more to do with uncertain, undesired outcomes, volatility is not a flawless representation of risk as changes in volatility can be a result of desired outcomes as well. As noted by Poon (2005), for a holder of a stock or other security, increase in volatility due to surging stock price is one example where high volatility is favoured. However, as noted initially by Black (1976), volatility tends to have a negative relationship with returns, meaning that downward movements in the price of the underlying tend to be more pronounced that upward movements, resulting in higher volatility. This effect is commonly referred to as leverage effect. As volatility is an inherent feature in the markets, sources of volatility are ambiguous, often related to disclosed information, trading pressure, returns and spillovers of volatility between markets (Hull, 2017; Gutierrez, Martinez & Tse, 2009). As noted by Müller (1993), impact of new information can also flow to markets in a non-synchronous manner due to heterogeneity of market participants and nature of their investing philosophy. Heterogenous markets hypothesized by Müller (1993) will be revisited shortly.

### 2.1.1. Realized volatility

Historical, or realized, volatility is often presented as the standard deviation of e.g. asset returns, interest rates and commodity prices in the domain of finance. It is often necessary to determine volatility for different periods, say one-week or one-month ahead. However, for comparability, it is common to present volatility as an annualized figure. Common assumption regarding price development is that prices follow geometric Brownian motion (Hull, 2017). If this holds true, volatility estimated for any period can be used to derive the estimated standard deviation for any period $T$. Naturally, longer the time horizon in question, higher the uncertainty of outcome and thus the estimated standard deviation will be. (Brooks, 2005; Hull, 2017)

As true level of volatility is latent even after it has realized, a proxy variable has to be used to represent volatility. As noted before, it is common to use standard deviation of returns or squared returns to estimate volatility, for which both low- and high-frequency data can be used. Squared daily returns from closing price to closing price have been used extensively, however, as noted by for instance by Kang and Maysami (2008), daily volatility estimates calculated from high-frequency intraday data seem to be more robust and consistent estimators of the true level of volatility. In the absence of intraday data, other proxy variables have been formulated, for instance a drift-independent volatility estimator by Yang and Zhang (2000), which will be used in this thesis as a proxy for latent volatility. Yang-Zhang estimator and other possible volatility proxies are presented in chapter 2.1.3.

### 2.1.2. Implied volatility and volatility smile

In addition to realized volatility, option markets can be used to obtain forward-looking option implied volatilities, commonly viewed as the consensus estimate of market expectations regarding average volatility of the underlying asset, during the lifetime of the option (Hull, 2017). As noted earlier, implied volatility and Black & Scholes-option pricing model are intimately connected. As latent volatility in B&S is the single non-observable variable in the market, the model can be used to infer B&S volatilities from option prices. As noted by Lee (2004), it is common practice in some option markets to quote options not by price but their implied volatility. In this thesis, B&S implied volatilities of S&P 500 options obtained from Refinitiv Eikon are used.

One of the biggest breakthroughs in the field of finance during the last century was Black-Scholes (1973) or Black-Scholes-Merton option pricing model, changing the way how practitioners price and hedge option contracts. Even though the model is seldomly credited to Fischer Black and Myron Scholes, as the name suggests, it is Robert Merton who developed a more general approach to deriving the model. Instead of relying on CAPM, as Black and Scholes did, Merton constructs a portfolio including a long (short) position in an option and a short (long) position in the underlying, non-dividend paying stock. Thus, Merton argues, the portfolio is risk-free, at least in the short-term, and should consequently yield a risk-free return. Merton's approach to option valuation is hence risk-neutral; expected return of either the stock or the option is not required and thus the method is indifferent to investors' risk preference. Due to the risk-neutrality Merton's approach is more general than that provided by Black and Scholes. (Hull, 2017)

As often is the case with theoretical finance and economics, the model is subject to quite strict and doctrinaire assumptions. As noted by Hull (2017), in the B&S world, arbitrage opportunities are non-existent and no costs or taxes are paid for transactions and profits. Trading is continuous and investors can lend and borrow at constant risk-free rate during the life of the option. Furthermore, stocks are perfectly divisible and pay no dividends during the life of the option, and importantly, their prices follow geometric Brownian motion.

Assumption of Brownian price process results in a lognormal distribution of stock prices with a constant expected return $\mu$ and constant variance $\sigma$.

Should the assumptions hold, the implied volatility for all options with the same underlying and maturity are equal. However, it has been long evident that rather than having equal volatilities, options with different strike prices, i.e. moneyness, differ in implied volatility and form a pattern resembling a smile or a smirk when implied volatility is plotted against moneyness. As noted by Pena, Rubio and Serna (1999), B&S is often thought to misprice in-the-money and out-of-the-money options. A common source of said mispricing is thought to be related to non-constant volatility in the real world, which often results in non-lognormal price distribution (Hull, 2017; Dumas, Fleming & Whaley, 1998). Financial data often tends to be non-normally distributed, with time series data including multiple stylized features, such as fat-tailed distributions or leptokurtosis, leverage effects and volatility clustering (Brooks, 2005). A rough visualization of commonly observed volatility smile and a "volatility plane" in line with B&S assumptions is provided in Figure 1.

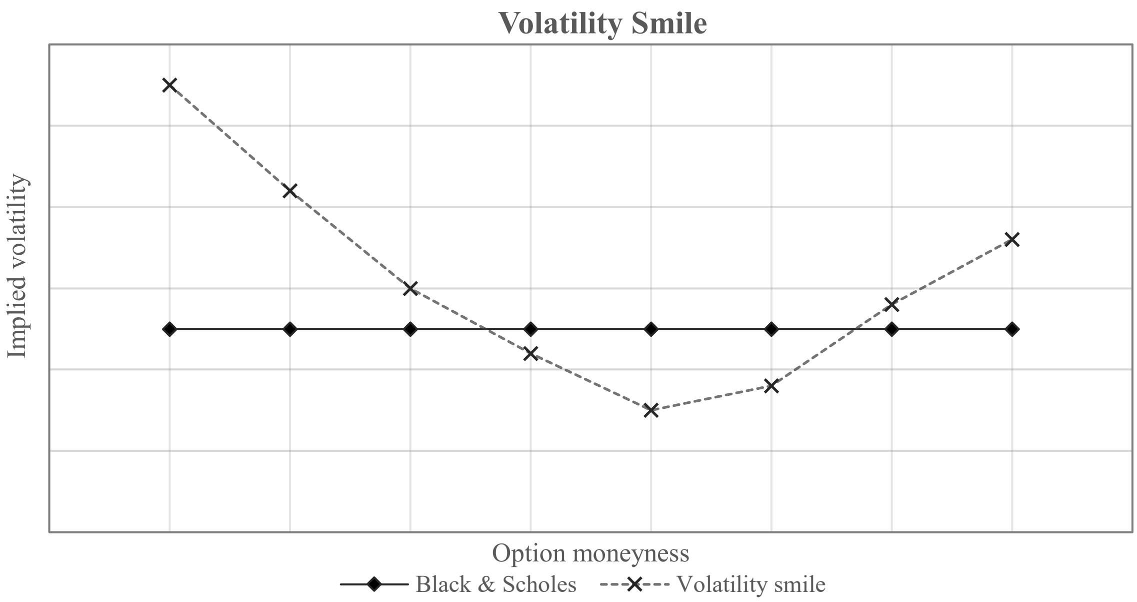



*Figure 1: Example illustration of the implied volatility smile*

Impact of higher moments of the underlying price distribution, that is, skewness and kurtosis, on the shape of the volatility smile has been researched, yet no clear-cut consensus has been reached. Among others, Bakshi, Kapadia and Madan (2003) and Backus, Foresi and Li (2004) have studied the relationship of underlying physical distribution and the shape of the volatility smile and have provided empirical evidence of this relationship. In equity markets, the physical distribution tends to exhibit excess kurtosis and negative skewness, due to for instance leverage effect and price jumps (Yan, 2011), thus leading to a convex, left-leaning smile or smirk. Corrado and Su (1997) note that for a left-skewed distribution, deep out-of-the-money put options tend to exhibit higher implied volatility than deep-out-of-money call options, as the probability of downward movement is higher with a left-skewed distribution than what is expected with normally distributed returns. In currency markets, the smile tends to be visibly more symmetrical than in the equity markets due to the absence of skewness but in the presence of excess kurtosis (Backus et al., 2004).

Other explanations for the existence of the smile have been proposed Bollen and Whaley (2004), Ederington and Guan (2005) and Pena et al. (1999). Net buying pressure hypothesis by Bollen and Whaley (2004) is based on the notion that order imbalance for index put and call options due to hedging pressure tends to drive a wedge between theoretical option price and observed option price. As market makers tend to be writers of out-of-the-money put options, costs of hedging their unbalanced positions are substantial and thus market makers quote bid prices higher for e.g. out-of-the-money put options (Bollen & Whaley, 2004; Ederington & Guan, 2005). Also Pena et al. (1999) emphasize the impact of transaction costs proxied by bid-ask spread as the main factor driving the smile, although they do underline the theoretical justification for the smile in the form of leptokurtic return distribution and non-constant volatility. Pena et al. (1999) note that high volatility periods tend to be accompanied by with less convex smile.

Implied volatility smile in this thesis is depicted with quadratic implied volatility function, or, deterministic volatility function, DVF. The methodology is well-represented in the literature (Dupire, 1994; Dumas et al., 1998; Bakshi et al., 2003; Backus et al., 2004; Bollen and Whaley, 2004; Buyn and Kim, 2013; Wong and Heaney, 2017). In DVF framework, implied volatility is regressed against option moneyness and its square, due to the parabolic nature of the smile. In some settings also time-to-maturity of the option and its square have been added as independent variables in the quadratic function. However, Dumas et al. (1998) note that this type of model is prone to substantial overfitting and does not improve the option pricing ability of a valuation model. Hence, only moneyness and squared moneyness are used as independent variables. Quadratic volatility function in this thesis is formulated as

$$\sigma_{IV,t} = \alpha_0 + \beta_1 M_t + \beta_2 M_t^2 + \omega_t \qquad (1)$$

where $\sigma_{IV,t}$ is implied volatility and $M_t$ is the option moneyness at time t, calculated as the ratio of option strike price and price of the underlying. First and second derivatives of implied volatility function are then used as proxies for the skewness and kurtosis of the underlying risk-neutral distribution, a method based on the work of Backus et al. (2004) and adopted by at least Wong and Heaney (2017). Further discussion regarding implied volatility functions and used variables are provided in Chapters 3 and 4.

### 2.1.3. Proxying volatility

As discussed earlier, volatility, or the conditional variance, is not observable even after the volatility is realized. Thus, a decision has to be made on which variable to use as a proxy for the true level of latent volatility; this proxy variable can then be interpreted as an ex-post approximation of the true level of realized variance during the period. As noted by at least Patton (2011), choice of volatility proxy is of utmost importance as it can greatly impact the performance evaluation of out-of-sample volatility forecasts when used as a benchmark. Biased volatility proxies used as a benchmark can then lead to inadequate and inconsistent model rankings when forecasting accuracy is compared. In addition, biased proxies can result in loss of explanatory power when used as independent variables in volatility forecasting models (Patton, 2011). Despite its importance, no clear consensus has been reached regarding the practice of estimating or calculating a volatility proxy.

After technological advances and increased availability of high-frequency data, one of the most common volatility proxies has been cumulative squares of intraday returns, argued by Kang and Maysami (2008) to be robust and unbiased estimates of the true latent volatility. While the methodology for intraday variance was first developed for the purpose of modelling intraday volatility, it has been widely used in literature for longer periods with less frequent sampling (Blair, Poon & Taylor, 2001; Kang and Park 2008; Corsi 2009). Volatility measures based on high-frequency data still contain some inherent shortcomings and can be biased estimates of the true volatility. A common concern regarding high-frequency volatility estimates is related to market microstructure, as sudden moves of bid-ask spreads can lead to increased high-frequency volatility and once cumulated across the day, can lead to severely overestimated volatilities. (Alizadeh, Brandt & Diebold, 2002) Regarding this thesis, no intraday data besides opening and closing prices as well as daily high and low prices were available for S&P 500 index, so intraday estimates of volatility cannot be considered.

In the absence of high-frequency data, squared returns are considered as a volatility proxy. While the method of proxying volatility with squared returns might seem counterintuitive at first, reasoning behind the squared returns approach is deeply rooted in common volatility forecasting practices. For returns, the variance is calculated as in (2), with long-term mean return often assumed to be zero, as this characterizes the long-term mean more accurately than the sample mean, which can be highly biased measure particularly in smaller samples. (Figlewski, 1997) When the long-term mean is set to zero, it is evident from the formulation of the variance function that volatility is only dependent on squared returns, hence its usage as a proxy for volatility. Andersen and Bollerslev (1998) argue that squared daily returns is a highly noisy and biased volatility proxy and according to Patton (2011) can lead to severely inadequate and inconsistent forecast evaluations. Variance of returns and close-to-close volatility estimate (3) are formulated as follows:

$$\sigma_t^2 = \frac{1}{n-1} \sum_{t=1}^{n} r^2 \tag{2}$$

$$\sigma_{close-to-close,t} = \sqrt{\frac{N}{n-2} \sum_{i=1}^{n-1} {r_i}^2} \tag{3}$$

where $r_i$ represents the logarithmic daily return for period $i$, $n$ is the total number of periods and $N$ is the number of trading days in a year, with 252 used per common practice. As sample means in (2) and (3) are assumed as zeros, these functions only include squared returns.

To combat bias and noise in volatility proxy estimated by squared returns, several alternatives have been developed and applied widely in the existing literature concerning volatility forecasting. Most notable and widely-spread work has been done by Parkinson

(1980), Garman and Klass (1980), Rogers and Satchell (1991) and Yang and Zhang (2000). These alternative proxies utilize intraday price ranges, such as open-close and high-low ranges, and thus are more able to depict the intraday variance and thus give more efficient approximations of realized volatility than plain squared returns. Furthermore, range-based estimates tend to be less biased regarding bid-ask spread movements and other market microstructure considerations than high-frequency volatility estimates (Alizadeh et al., 2002). Extensive availability of opening, closing, high and low prices for several assets over long periods promotes the utilization of range-based volatility proxies.

Usage of intraday price ranges in volatility estimation was pioneered Parkinson (1980), who introduced an estimator based on high and low prices during the trading day. In the proxy by Parkinson (1980), realized volatility is estimated as a logarithm of high-low ratio during the day, and as argued by Parkinson (1980), it provides a better proxy than squared returns for realized volatility, if the logarithmic price process is assumed to follow a zero-drift Brownian motion, that is, the expected return for the period is zero. High-low range proxy contains further information regarding intraday price fluctuations than squared returns, and is thus able to provide more accurate estimation of volatility. Obvious example of the superiority of high-low range over squared returns is a highly volatile trading day but where market closes close to opening levels; close-to-close estimator would greatly underestimate the true volatility when compared against high-low estimator. Inspired by Parkinson (1980), Garman and Klass (1980) extended the high-low proxy by including opening and closing prices to the estimator, similarly assumed to follow a Brownian motion with zero expected return. High-low (4) and Garman-Klass (5) estimators are formulated as:

$$\sigma_{P,t} = \sqrt{\frac{N}{4n \times \ln(2)} \sum_{i=1}^{n} \ln\left(\frac{H_t}{L_t}\right)^2} \tag{4}$$

$$\sigma_{GK,t} = \sqrt{\frac{N}{n} \sum \left[\frac{1}{2}\left(\ln\left(\frac{H_t}{L_t}\right)\right)^2 - (2\ln(2) - 1)\left(\ln\left(\frac{C_t}{O_t}\right)^2\right]} \tag{5}$$

where $H_t$ and $L_t$ denote the high and low prices and $C_t$ and $O_t$ the closing and opening prices of the day $t$. Volatility estimators by Parkinson (1980) and Garman and Klass (1980) have few noteworthy shortcomings, as noted by Kang and Maysami (2008). First, as they are built on the assumption of Brownian price process with expected return equal to zero, they might lead to biased estimates for asset with non-zero returns during the estimation period. Second, continuous trading is assumed in their model, leading to volatility estimates lower than the realized latent volatility to some extent. Third, neither of the estimator include overnight jumps between yesterday's closing and today's opening price, possibly leading to further underestimations of volatility. To tackle some of these issues, Rogers and Satchell (1991) introduced a volatility estimator for assets with non-zero mean returns. Nevertheless, their estimator for latent volatility does not take into account overnight jumps, leading to possible under estimations of realized volatility. Estimator developed by Rogers and Satchell (1991) is formulated as:

$$\sigma_{RS,t} = \sqrt{\frac{N}{n}\sum\left[\ln\left(\frac{H_t}{C_t}\right) \times \ln\left(\frac{H_t}{O_t}\right) + \ln\left(\frac{L_t}{C_t}\right) \times \ln\left(\frac{L_t}{O_t}\right)\right]} \tag{6}$$

In the presence of overnight jumps, an estimator introduced by Yang and Zhang (2000) can be used to obtain the most accurate range-based volatility estimates, as other methods ignore the impact of differences between yesterday's closing and today's opening prices, thus underestimating the realized volatility (Shu & Zhang, 2006). Yang-Zhang estimator is drift-independent and takes into account overnight volatility, thus tackling the issues inherent in earlier volatility estimators. Yang-Zhang estimator is formulated as:

$$\sigma_{YZ,t} = \sqrt{N}\sqrt{\sigma^2_{overnight,t} + k\sigma^2_{open-close,t} + (1-k)\sigma^2_{RS,t}} \tag{7}$$

where

$$\sigma^2_{overnight,t} = \frac{1}{n-1}\sum_{t=1}^{n}\left[\ln\left(\frac{O_t}{C_{t-1}}\right) - \overline{\ln\left(\frac{O_t}{C_{t-1}}\right)}\right]^2$$

$$\sigma^2_{open-close,t} = \frac{1}{n-1}\sum_{t=1}^{n}\left[\ln\left(\frac{C_t}{O_t}\right) - \overline{\ln\left(\frac{C_t}{O_t}\right)}\right]^2$$

$$k = \frac{0.34}{1.34 + \frac{n+1}{n-1}}$$

Due to its ability to accurately estimate latent volatility, Yang-Zhang estimator for realized volatility is used in this thesis to calculate volatility proxies for daily, weekly and monthly periods.

## 2.2. HAR-RV

As volatility and particularly its forecasting has remained as one of the most essential concepts in finance over several decades, serious effort has been put into developing models able to reproduce common features in asset returns, such as volatility clustering and leptokurtic returns. Popularity has been gained by models of GARCH family, initially introduced by Taylor (1986) and Bollerslev (1986) and later developed for instance by Nelson (1991) As GARCH models conditional variance using lags of residual and lags of conditional variance itself, it is able to capture volatility clustering and thus replicate some leptokurtosis. Despite its popularity, GARCH models have some shortcomings.

Empirical evidence of scaling and multi-scaling, i.e. repetition of patterns over horizons differing in length, in returns has been provided by Di Matteo (2007) and Corsi (2003; 2009), among others. Scaling can thus be thought to imply that volatility arrives to the markets as waves with distinct time scales, requiring models to be able to reproduce long-memory characteristics. However, the feature cannot be properly captured by GARCH and short-memory stochastic volatility models, as their volatility tends to resemble white noise once aggregated over longer horizons. In addition, true long-memory models such as ARFIMA can be cumbersome to fit, and their economic parametrical interpretation tends to be near-impossible. (Corsi, 2009) Thus, Corsi (2003; 2009) has proposed a parsimonious time-series method for modelling conditional volatility with partial components of volatility realized over different horizons. In general, motivations for the model are the ability of a simple model to reproduce long-memory and other common characteristics observed in financial data and the economic observation of volatility cascades as a by-product of heterogeneity in markets, discussed shortly.

Heterogenous Autoregressive model of Realized Volatility by Corsi (2003; 2009), HAR-RV, models future volatility through partial volatility components estimated over periods differing in length. As such, separate components then function almost as autoregressive processes of order 1 estimated over distinct time horizons, hence the model is "heterogeneously autoregressive." Partial volatilities in HAR-RV represent distinct elements of volatility generated by heterogenous market components. HAR-RV by Corsi (2003; 2009) is formulated as:

$$\sigma_{t,T} = \alpha_0 + \beta_1 \sigma_{t-1,d} + \beta_2 \sigma_{t-5,w} + \beta_3 \sigma_{t-21,m} + \omega_t \tag{8}$$

where $\sigma$ denotes realized volatility for future period $T$ forecasted at time $t$, and $d$, $w$ and $m$ represent daily, weekly and monthly periods for which realized volatility is estimated. Due to its additive nature, the model is highly parsimonious and offers clear economic interpretation. Furthermore, structure of the model enables easy adoption of additional regressors, a feature serving well the purpose of this thesis, and also utilized in previous research (Andersen, Bollerslev and Diebold, 2007; Buyn and Kim, 2013). While true long-memory properties are absent in the model due to its simplicity, simulation results by Corsi (2009) show that the model is able to reproduce critical empirical features of financial returns, such as volatility clustering and fat tails. Furthermore, empirical results in Corsi (2009) show great out-of-sample forecasting performance as the model constantly outperforms short-memory models and is equal to more parsimonious long-memory ARFIMA-model. In addition, the inclusion of Yang-Zhang volatilities as proxies for partial volatility is straight-forward and thus volatility can be modelled using more accurate proxies than squared returns.

From economic perspective, HAR-RV is inspired by heterogenous market hypothesis by Müller (1993) recognizing the heterogeneity among market participants. Müller (1993) argues that agents in markets differ in proficiency, risk appetite and institutional constraints as well as in trading frequency and geographic location. As noted by Corsi (2003; 2009), HAR-RV is particularly based on temporal heterogeneity. Time horizons of agents range from frequent intraday dealing and market making to multi-year investment mandates of pension funds, and thus market participants process and react to information distinctively, producing components of volatility dissimilar in time scale. Empirical observations of financial data have shown that long-term volatility tends to have greater impact on short-term volatility than vice versa (Müller, 1997; Lynch and Zumbach, 2003). In HAR-RV, structure of unique volatility components is simplified to cover daily, weekly and monthly periods, representing daily traders, medium-term agents with weekly rebalancing, for instance, and long-term investors with carefully processed trading decisions. (Corsi, 2009) HAR-RV is thus preferred as a model not only by its parsimonious structure and great performance shown by Corsi (2003; 2009), but also due to economic rationale and clear economic interpretation.

### 2.2.1. Evaluation of forecasts

Once forecasts for daily, weekly and monthly volatility are done, model performance is evaluated by assessing forecasting accuracy. Previous literature on forecasting volatility includes an extensive selection of distinct evaluation measures, often comparing differences in forecasted and realized volatility for the period. As noted by Patton (2011) among others, complications in performance evaluation can be encountered as these evaluations are done against imperfect proxies of true volatility. Thus, it is important to use robust evaluation measures when assessing forecast errors to ensure truthful ranking of the models. In this thesis, root mean squared error (RMSE) and mean absolute error (MAE) are used to assess out-of-sample forecasting performance.

Patton (2011) provides evidence that mean squared error, MSE, is a feasible choice for error measure when using imperfect volatility proxies, since it is robust to noise in these and in general ranks models consistently even when different proxies for latent volatility are used. According to Patton (2011), consistent relative rankings of models are obtained when using MSE with a certain proxy or when true latent volatility would be used. Mean squared error is chosen in this thesis due to its robustness; however, to ease the comparison of forecast errors and average volatility during the period, RMSE is preferred over plain MSE as taking the square root of MSE alters the measure to be more informative of error size. RMSE in this thesis is formulated as:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(\sigma_F - \sigma_t)^2}{N}} \tag{9}$$

where $\sigma_F$ denotes the forecasted volatility for time t, $\sigma_t$ represents the actual realized volatility at time t and $N$ is the total number of forecasted values. In addition to RMSE, mean absolute error is used as a secondary evaluation measure of the models. As forecast errors are squared when calculating RMSE, the measure is inherently prone to outliers in volatility and thus penalises model performance during high-volatility periods, even when no decline is seen in the actual forecasting accuracy. Thus, MAE is deployed as an additional evaluation metric to improve the robustness of results. MAE is formulated as:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\sigma_F - \sigma_t| \tag{10}$$

# 3. Literature Review

This chapter provides an extensive review of relevant literature considering volatility forecasting. First, earlier research regarding the HAR-RV model (Corsi, 2003; 2009) used in the thesis is introduced. Then, quite contradictory set of evidence of whether volatility forecasting should be done using historical or implied volatility is provided. Finally, earlier research is introduced regarding information content in the implied volatility skew as well as proxying the shape of the skew in volatility forecasting. Contradicting results of implied versus historical volatility and the methods in depicting the skew provide the motivation for the methodology used in this thesis.

## 3.1. Previous literature on HAR-RV

Volatility forecasts in this thesis are done with Heterogenous Autoregressive model for Realized Volatility, HAR-RV, first introduced in Corsi (2003) and Corsi (2009). The model has gained popularity due to its parsimonious nature, as model extension via additional regressors is straightforward, as well as due to its ability to reproduce stylized properties observed in financial data, such as leptokurtosis and clustering of volatility. Furthermore, despite its lack of true long-memory elements the model is able to reproduce these features. Corsi (2003; 2009) provides evidence of the model's adequacy in volatility modelling in equity, currency and interest rate markets by simulating realized volatility of S&P 500 futures, USD/CHF exchange rate and 30-year U.S. Treasury futures over a period of 14 years, from 1989 to 2003. Corsi (2009) provides simulation results including daily returns and their probability density functions and realized volatilities and autocorrelation present in simulated volatility, as well as their realized counterparts from the actual returns. In addition, evidence of forecasting ability is provided by Corsi (2009) for one-day, one-week and two-weeks ahead forecasts of volatility, with daily model re-estimations over a rolling period of 1000 observations. Corsi (2009) used AR(1), AR(3) and ARFIMA models as benchmarks, evaluates forecasting performance with RMSE, MAE and $R^2$ values. Corsi (2009) shows evidence of overperformance of HAR-RV relative to short-memory models and equal performance with more parsimonious long-memory ARFIMA model. Simulation results by Corsi (2009) verify that even despite HAR-RV includes no proper long-memory characteristics, it is able to reproduce stylized facts observed in financial data.

As noted, the model has been adopted and extended frequently in previous literature due to its parsimonious nature and ability to replicate common characteristic in financial data. Andersen, Bollerslev & Diebold (2007) extend the HAR-RV-model by Corsi (2003) by including a jump component to differentiate impacts of sporadic jumps from the continuous sample path. They demonstrate increases in volatility forecasts using high-frequency intraday data in currency, equity and rate markets, namely by studying Deutsche Mark/USD exchange rate, S&P 500 index returns and thirty-year Treasury yields. Andersen et al. (2007) provide evidence of the non-persistent and unpredictable nature of the jump component and conclude that only the continuous sample path component increases predictive power in volatility forecasts.

Similar methodology has been employed by Corsi, Pirino and Reno (2008) who also study the impact of jumps in volatility forecasting using HAR-RV model enhanced with additional regressors to account for jumps. They use data containing almost 15 years of high-frequency data for S&P 500 and thirty-year Treasury futures, as well as nearly five years of high-frequency data for six individual stocks. They employ similar extension of HAR-RV introduced by Andersen et al. (2007) and find evidence that jumps have a positive impact on volatility, with this relationship being statistically significant most of the time. Authors in both Andersen et al. (2007) and Corsi et al. (2008) underline the usefulness of HAR-RV in parsimonious volatility forecasting with long-memory characteristics.

Audrino and Hu (2016) extend the HAR-RV framework to study impact of both jumps and leverage effect on volatility forecasting. Dataset used in Audrino and Hu (2016) contains tick-by-tick high-frequency price data of S&P 500 futures with extremely long period spanning from 1982 to 2010. They separate upside and downside risks of realized volatility as well as continuous and irregular components of the leverage effect in their methodology and show evidence that adverse price jumps tend to be more informative of future volatility than upside moves. Furthermore, they note that continuous leverage component tends to be more lasting than the irregular one. Asymmetries in volatility have been also studied by McAleer and Medeiros (2007) who propose an extension for HAR-RV by enhancing the baseline model by including variables representing multiple regimes of sign and asymmetries. Their dataset consists of tick-by-tick high-frequency data for 16 companies in the Dow Jones Industrial Average index with a timespan from 1994 to 2003 and find strong evidence the support the persistent nature of both sign and size asymmetries. Both Audrino and Hu (2016) and McAleer and Medeiros (2007) opt for the HAR-RV model to be extended in their study due to its parsimonious nature and ability to reproduce commonalities in financial data together with long-memory characteristics.

Previous extensions of HAR-RV by Corsi (2003; 2009) provide evidence of the models ability to perform in volatility forecasting as well as its usability as a baseline model for different extensions, due to its parsimonious nature. Several of the extensions have focused on including components of jump risk or asymmetries of volatility, but inclusion of volatility smile components to HAR-RV has also been introduced, namely by Byun and Kim (2013) whose research will be presented shortly. As such, this thesis somewhat follows their methodology but over a more recent period.

## 3.2. Forecasting volatility

An extensive survey conducted by Poon and Granger (2003) provides some evidence in favour of using implied rather than historical volatility as an predictor for future volatility. Poon and Granger (2003) survey 93 published papers regarding volatility forecasting and conclude that implied volatility models tend to perform the best; performance of different historical models tend to perform more or less equally, ranging from exponentially weighted moving average models to ARMA and GARCH-type of models as well as to stochastic volatility models. While Poon and Granger (2003) suggest that implied volatility models seem to be preferable than historical volatility models, they note that papers surveyed rarely provide extensive comparisons against different types of models, leaving the question of what model to prefer often unanswered. In addition, no proper conclusion can be made about which models are preferred for different forecast horizons; the most common forecast horizon tends to be one day.

First contributions in volatility forecasting with option implied volatility were made by Latané and Rendleman (1976), Chiras and Manaster (1978) and Beckers (1981). Latané and Rendleman (1976) studied 24 large publicly traded U.S. companies which had options trading on the Chicago Board Options Exchange (CBOE) for a 39-week period between October 1973 and June 1974. By using Black-Scholes option pricing model, they derived weighted average standard deviations implied by all suitable traded options and conducted static cross-sectional tests against realized standard deviations of underlying stock prices. Latané and Rendleman (1976) concluded that weighted average standard deviations derived with Black-Scholes were significantly more correlated with realized stock price standard deviations than deviations derived from historical price data.

Both Chiras and Manaster (1978) and Beckers (1981) build on the framework provided Latané and Rendleman (1976) and study individual stock options traded on the CBOE. Beckers (1981) found evidence that particularly at-the-money options provided most accurate forecasts of future variance, thus specifying the conclusion of Latané and Rendleman (1976), who did not make distinctions between options differing in moneyness. Results provided by Chiras and Manaster (1978), on the other hand, suggest that at the time it was plausible to build an riskless trading strategy capable of producing abnormal returns, by using the informational content in implied standard deviations derived by Black-Scholes.

When studying the informational content in S&P 100 index options, Day and Lewis (1992) find mixed evidence of whether implied volatility has incremental predictive power over historical volatility. They use weekly returns calculated with daily closing prices of S&P 100 index as well as daily implied volatilities for index options, and forecast future volatility with GARCH- and EGARCH models. Day and Lewis (1992) then use option implied volatilities as exogenous regressors in these models and compare the predictions. Their in-sample results suggest that implied volatility might contain additional information regarding future volatility not present in conditional variances obtained from GARCH-models. However, they find no proper evidence that either GARCH conditional variances or implied volatility completely represent the actual in-sample conditional variance in stock markets.

Building on the results of Day and Lewis (1992), Lamoureux and Lastrapes (1993) study ten individual stocks and their options traded on the CBOE, using daily price data between 1982 and 1984. To tackle the issue of constant volatility assumption in Black-and-Scholes, implied volatilities in Lamoureux and Lastrapes (1993) are derived with a class of stochastic volatility models provided by Hull and White (1987). Their results have two implications. First, implied volatilities derived with Hull-White model tend to predict lower future volatility when compared to realized volatility. Second, volatility forecasts with past returns tend to contain relevant information regarding future variance not present in market forecasts, i.e. implied volatilities, even though they may not provide more accurate predictions per se.

Poor results regarding information content of implied volatility are provided by Canina and Figlewski (1993), who study S&P100 index options between 1983-1987. They use daily index and option price data but exclude options with less than 7 days and more than 127 days until maturity. In addition, they exclude options with strikes 20 percent higher or lower than the underlying index. When assessing the predictive power of models utilizing either implied volatility or historical volatility, Canina and Figlewski (1993) conclude that implied volatility is a poor predictor of future realized volatility, as during their research period, implied volatility exhibits virtually no correlation with future volatility. Moreover, information contained in recent observed volatility is not incorporated in implied volatility in their sample.

Contradictory evidence regarding the usefulness of implied volatility as forecast of future volatility is provided Christensen and Prabhala (1998). When studying the relationship between implied and ex post realized volatility using S&P 100 index and index options, Christensen and Prabhala (1998) find evidence that implied volatility outperforms historical volatility when forecasting volatility in equity markets. In contrast to earlier studies, they use a longer period of 11 years and monthly sampling, resulting in non-overlapping data, a common concern in several earlier studies, including Day and Lewis (1992) and Lamoureux and Lastrapes (1993). Their results suggest that implied volatility subsumes information of historical volatility and that implied volatility is an efficient predictor of future volatility. In addition, Christensen and Prabhala (1998) conclude that historical volatility has far less explanatory power than implied volatility, inconsistent with Canina and Figlewski (1993), and with some of their model specifications, it has no additional information not already included in implied volatility.

Blair, Poon and Taylor (2001) study the forecasting performance of models using historical volatility as well as implied volatility. Their dataset includes both daily closing prices and intraday prices with 5-minute-intervals between 1987 and 1998 for S&P 100 index, which are used to calculate two different volatility proxies. In addition, to depict implied volatility, they use VIX index which tracks implied volatility of several S&P 500 index options. Blair et al. (2001) find evidence that VIX provides most accurate forecasts for all horizons, ranging from one-day to 20-day volatility. Their results suggest that for one-day volatility forecasts, ARCH-forecasts with daily returns might contain some additional predictive power on top of VIX forecasts; however, no significant evidence of this is found.

Correspondingly to Blair et al. (2001), Giot (2005) uses different implied volatility indices, VIX, VXN and VXO, during a period from 1994 to 2003. To study the information content of volatility indices, Giot (2005) utilizes them as a volatility input in Value at Risk-models. Performance of VaR-models using different volatility indices is then compared against the performance of VaR-models with volatilities modelled with GARCH-models. Results in Giot (2005) suggest that implied volatility indices seem to include all significant information regarding future volatility and thus subsumes the information of GARCH volatilities. Giot (2005) provides evidence that these results are robust for different types of markets, including bull market with high and low volatility regimes as well as bear market with high volatility.

In more recent research, Kambouridis, McMillan and Tsakou (2016), who study volatilities of different U.S. stock indices, S&P 500, Dow Jones Industrial Average and Nasdaq 100 as well as their respective implied volatility indices VIX, VXD and VXN between 2001 and 2013. Kambouridis et al. (2016) compare the volatility forecasting performance of several types models based on historical data, such as ARMA and GARCH-type models, as well as based on implied volatility. Their results suggest that plain models using only implied volatility as a predictor perform worse than alternate models such as GARCH. However, best forecasting performance for all datasets is delivered by model combining implied volatility with asymmetric GARCH-model. Hence, results in Kambouridis et al. (2016) advocate for the usage of model combining both implied and historical volatility.

## 3.3. On the shape of implied volatility smile

Extensive research has been conducted in effort to explain the shape of the volatility smile, or implied volatility function, IVF, yet no certain conclusions have been made. For instance, Dupire (1994) and Derman and Kani (1994) have created an implied binomial tree framework able to depict the volatility structure of options differing in moneyness with in-sample data. However, empirical tests by Dumas, Fleming and Whaley (1998) provide evidence that parameters of binomial tree in Dupire (1994) and Derman and Kani (1994) are highly unstable through time and cannot properly explain the over-time variation in the shape of the implied volatility function. At least Bakshi, Cao and Chen (1997) and Chernov, Gallant, Ghysels and Tauchen (2003) find evidence that stochastic volatility models enhanced with jump component are able to replicate the shape of the IVF in-sample, Bakshi et al. (1997) for S&P 500 index options and Chernov et al. (2003) for Dow Jones Industrial Average. However, as noted by Bollen and Whaley (2004), parameters in their models are (1) unstable and (2) highly different than those calculated from actual returns, and thus they can explain the shape of the IVF only partially. To overcome these issues, Bollen and Whaley (2004) as well as Ederington and Guan (2005) theorize that trading pressure or net buying pressure could explain the shape of the implied volatility function.

Net buying pressure hypothesis is provided in Bollen and Whaley (2004) to explain the changes in the shape of the implied volatility smile. They use daily data of trades and quotes S&P 500 index options and 20 most traded individual stock options traded on CBOE between 1988 and 2000. Difference between buyer-motivated and seller-motivated traded contracts is then used as a proxy for net buying pressure. They find evidence that changes in implied volatilities for a certain option series are significantly related to net buying pressure. More specifically, for index option IVF the changes are driven by index puts, consistent with the notion that hedgers prefer index puts rather than calls for hedging their position, and for individual stock, the changes in shape of the IVF are driven by call options. In addition, Bollen and Whaley (2004) find evidence that IVF is significantly more negatively sloped for index options than individual stock option, a finding in line with Bakshi, Kapadia and Madan (2003).

Complimentary results in line with net buying pressure hypothesis is provided by Ederington and Guan (2005), who study S&P futures options traded on Chicago Mercantile Exchange between 1998 and 2003. Ederington and Guan (2005) argue that for the buying pressure hypothesis to hold, the changes in the shape of the implied volatility function do not need to have any predictive power regarding future volatility, as the changes can be driven by trading pressure in non-at the money-options due to hedging demand. They study the information content in implied volatilities of options differing in moneyness and find that only implied volatility of options near the smile's nadir are efficient and unbiased predictors of future volatility, with moneyness ranging from 1.02 to 1.04, i.e. option strike price 2 to 4 percent higher than the price of the underlying. When plotting the relative explanatory power of implied volatilities over option moneyness they find a pattern resembling a mirror image of volatility smile, which they term as "information frown."

Kang and Park (2008) study the effect of net buying pressure in Korean KOSPI 200 index option market. Similarly to Bollen and Whaley (2004), their dataset includes daily trades and quotes for KOSPI 200 index options, but they also utilize high-frequency intraday data with 5-minute intervals to study the dynamics between buying pressure, shape of the implied volatility function and price of the underlying index. Kang and Park (2008) find significant evidence that increase in buying pressure of call options precedes rising index levels five minutes later. An opposite effect is found regarding increasing buying pressure of put options. Their results suggest that the shape of implied volatility function is driven by informed directional traders rather than hedging pressure, a result in line with learning hypothesis provided in Bollen and Whaley (2004) as an alternative for buying pressure hypothesis: informed traders tend to trade first in the option markets, driving up the implied volatility of in the money- or out of the money-options, which is then "learned" by traders in equity markets.

Evidence of relationship between volatility skew and risk-neutral higher moments of the price distribution is found by Bakshi, Kapadia and Madan (2003), who study both the S&P 100 index and individual options. Their results suggest that (1) the more negatively skewed the physical distribution of future log prices of the underlying is, the more negative is the slope of implied volatility smile and (2) a higher risk neutral kurtosis leads to a flatter smile when the smile is left-skewed. The model provided in Bakshi et al. (2003) favours risk-neutral skewness over risk-neutral kurtosis in explaining the shape of the smile, which is modelled with a quadratic function of option moneyness, a common method in the literature (Dupire, 1994; Dumas et al., 1998). They also demonstrate that return distributions and thus the smile is more skewed for stock index than individual stocks, as idiosyncratic component of returns tend to be less skewed than the systemic component. Furthermore, researchers suggest that risk-neutral skewness regarding the index is a consequence of risk aversion and leptokurtic return distributions.

Additional evidence of the relationship between risk-neutral higher moments of the underlying future price distribution and shape of the volatility smile is provided by Backus, Foresi and Wu (2004), who study daily USD/CAD, USD/DEM and USD/JPY exchange rates and options between 1986 and 1996. Corresponding to earlier literature, they also depict the volatility smile as an quadratic function of option moneyness. Backus et al. (2004) derive risk-neutral skewness and risk-neutral kurtosis of the quadratic function with a method later adopted by at least Wang and Heaney (2017), and show evidence of the relationship between risk-neutral higher moments and the density of future log prices. Backus et al. (2004) conclude that in foreign exchange markets risk-neutral kurtosis is favoured over risk-neutral skewness in explaining the shape of the smile, as volatility smiles in exchange rate options tend to exhibit negligible to no skewness.

At least Dennis and Mayhew (2002) and Doran, Peterson and Tarrant (2007) find evidence of the relationship between ex ante skewness and ex post returns by using methodology for deriving risk-neutral skewness provided in Bakshi et al. (2003). Dennis and Mayhew (2002) study stock options traded on CBOE and find that implied skewness tends to be more negative for (1) high-beta stocks and (2) during times of higher market volatility, highlighting the importance of systemic risk in pricing individual stock options, in line with Bakshi et al. (2003). Results in Doran et al. (2007) regarding S&P 100 index options reinforce the notion that the smile does contain information of future returns, as they find evidence that the implied skewness can reveal the timing of market spikes and crashes with significant probability, at least in the short term with option maturities ranging from 10 to 30 days. However, Doran et al. (2007) note that while they find evidence of predictive power in the shape of the smile regarding future spikes and crashes, the magnitude of the prediction is not economically significant.

Ying, Zhang and Zhao (2010) find supplementary evidence of the relationship between future equity returns and slopes of individual equity options. Their study spans from 1996 to 2005 with daily data used regarding stock prices and option implied volatilities. Slope of the smile in Ying et al. (2010) is determined as the difference between implied volatilities of out-of-the-money put options and at-the-money call options, a method different than that suggested by e.g. Bakshi et al. (2003) and Backus et al. (2004). Ying et al. (2010) provide evidence that firms with the steepest smiles tend to underperform counterparties with least pronounced smiles, with a difference in annual returns of 10,6% on average. They find additional evidence that in general, companies with steepest smiles are the ones with the most adverse earnings shocks during the next quarter, concluding that option traders informed with negative news seem to trade OTM put options of these companies, and this information flows to equity market in a slow manner (Ying et al. 2010).

Yan (2011) studies the information content in the shape of the implied volatility smile regarding jump risk and finds evidence that the shape of the smile can be used as a proxy for jump risk in stock prices. Yan (2011) studies over 4000 individual stocks between 1996 and 2005, using monthly sampling for stock returns and option data, from which the slope of the smile is calculated. To depict the shape of the smile, Yan (2011) uses the difference between implied volatilities of at-the-money put and call options, a method deviant from the one used by at least Bakshi et al. (2003) and Doran et al. (2007), and from the one used in this thesis. Yan (2011) divides individual stocks in five different portfolios based on the slope of their smile, and provides evidence that the portfolio with lowest slope has the highest return during the period. This corresponds to his notion that shape of the smile is a valid proxy for average jump size, for which Yan (2011) provides empirical evidence.

Byun and Kim (2013) study S&P 500 options market and find that risk-neutral skewness significantly enhances accuracy of out-of-sample forecasts for daily, weekly and monthly volatility. Byun and Kim (2013) employ a forecasting model based on HAR-RV (Corsi 2009), and include implied volatility as well as risk-neutral higher moments in their model, an approach similar to one used in this thesis. Their results reaffirms that risk-neutral skewness, i.e. slope of the smile, contains additional information regarding future volatility not included in historical or implied volatility. They also compliment the work of Christoffersen, Elkamhi, Feunou and Jacobs (2010) in defining the connection between risk-neutral conditional variance and conditional moments of the physical distribution. Byun and Kim (2013) provide a theoretical relationship between the physical conditional variance and higher moments of risk-neutral distribution, thus justifying the usage of risk-neutral skewness and risk-neutral kurtosis as a proxy for the shape of the smile.

More recently, Wong and Heaney (2017) find evidence that including curvature of the skew to volatility forecast improves the one-month forecast accuracy in currency markets, using daily data for GBP/EUR, EUR/USD, AUD/USD and USD/JPY exchange rates. Forecast models used by Wong and Heaney (2017) are based on implied rather than historical volatility. As suggested in earlier literature, they depict the curvature with risk-neutral higher moments in line with Backus et al. (2004); however, they find little to no evidence that risk-neutral skewness or risk-neutral kurtosis dominates one another in improving forecasting accuracy in currency markets. Their results of predictive power in risk-neutral higher moments favour the learning hypothesis explaining the existence of the skew, as suggested by Kang and Park (2008).

Results in previous literature suggest than (1) implied volatility could be used to forecast future volatility and (2) the shape of the implied volatility smile seems to contain information regarding future volatility, at least to some extent. Thus, it seems reasonable to study whether these relationships have existed during the past years for S&P 500 index, a question this thesis tries to find answers to.

# 4. Data and Methodology

This chapter introduces used data and provides a comprehensive overview of the methodology used in deriving the implied volatility skew as well as variables depicting the shape of said skew. In addition, estimation of latent volatility as well as model training and forecasting procedure are presented. It is noteworthy to mention that while the two are closely related, implied volatility skew will be henceforth referenced as a smile or curve, to avoid confusions with risk-neutral skewness (of the physical distribution).

Data used in this thesis consists of daily opening, high, low and closing prices of S&P 500 stock index from 1.12.2017 to 17.9.2021, which are used to estimate latent volatility with Yang-Zhang-estimator. This period is chosen due to data availability as well as to include the crash period of 2020 instigated by Covid-19 crisis. In addition, to derive daily implied volatility curves, daily implied volatilities for S&P 500 index call options trading on Chicago Mercantile Exchange with tenors from one month to two years are used. Naturally, option strikes and maturity dates are also included in the data. Data is obtained via Refinitiv Eikon.

Table 1 presents relevant information of S&P500 index returns during the observation period, complemented by Figure 2 depicting the development of index price level and volatility in the form of daily returns. Noteworthy are the extreme minimum and maximum daily returns as well as kurtosis value, which are products of two-month period from March to April 2020 of exceptionally high volatility in the midst of Covid-19 crisis. Kurtosis of return distribution excluding this period is 3,94, a value much more comparable to that of normal distribution, 3. Return distribution shows also negative skewness with -0,68 overall and -0,85 excluding the above two-month period.

*Table 1: Distribution details of S&P500 daily returns*

| **S&P 500 Index Returns** | | | | | |
|---|---|---|---|---|---|
| Min | Max | Mean | Median | Kurtosis | Skewness |
| *-11,98 %* | *9,38 %* | *0,06 %* | *0,11 %* | *16,53* | *-0,68* |

To ensure robustness in implied volatility curve construction, significant amount of call options traded during the period are included in the option data, thus the original data set contains 1523 contracts for 2018, 1904 contracts for 2019, 2548 contracts for 2020 and 2239 contracts for 2021. All traded options ranging in moneyness from 0,875 to 1,125 are considered when constructing the curve, as IV tends to distort with extreme values of moneyness; motivation for moneyness range chosen comes from earlier literature (Bakshi & Kapadia, 2003; Doran et al., 2007). In addition, only options maturing within 120 to 10 trading days are included in curve construction. This decision is made based on three noteworthy considerations to avoid inconsistencies in the research. First, to ensure the consistency with earlier literature, as similar ranges have been employed by at least Bakshi et al. (2003), Doran et al. (2007) and Byun and Kim (2013). Second, options with time-to-maturity less than 10 trading days are excluded due to unreliable short-term IV as maturity approaches (Bakshi et al., 2003). Third, as S&P 500 index options mature quarterly, upper threshold of 120 days is used to ensure the ability to construct daily curves. Furthermore, to avoid calculation errors when deriving implied volatility curves, only the contracts with shorter maturity are considered in case of overlapping data. Thus, after sampling option data, yearly amounts of used contracts vary between 918 and 1040.

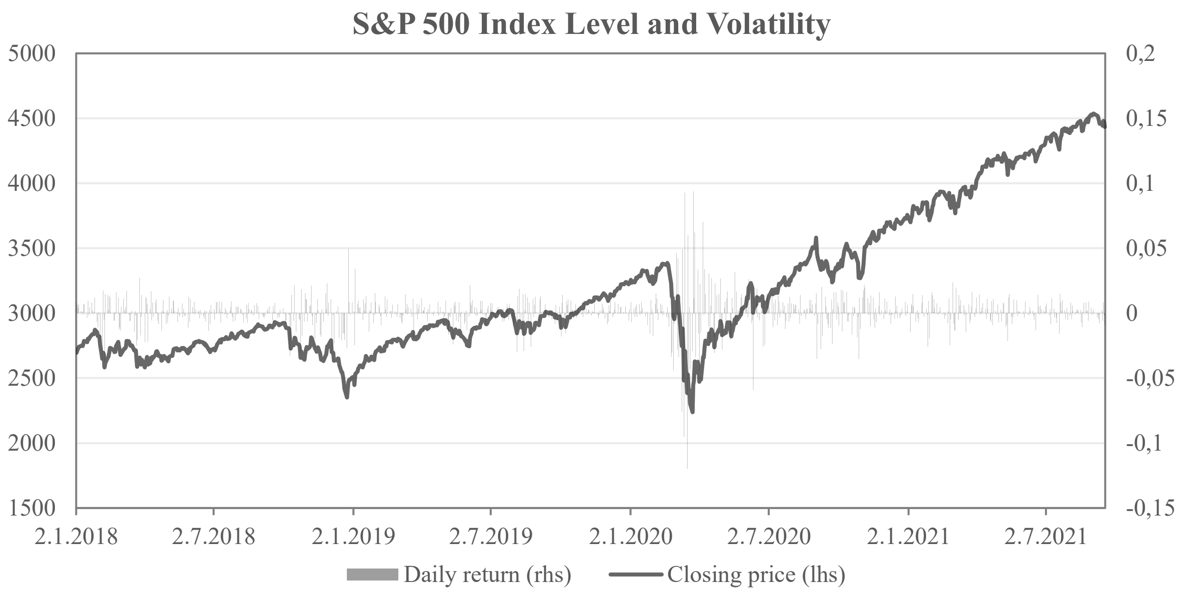



*Figure 2: S&P500 index level and volatility*

As IV tends to decrease in time, it is reasonable to account for time-to-maturity of the derived smile to ensure comparability of daily observations. This is done in line with Dumas et al. (1998) by adjusting the option moneyness with the square root of days left until maturity, as the dispersion rate of a stochastic process, such as the underlying index, with a time step $t$ tends to equal $\sqrt{t}$. Dumas et al. (1998) point out that the slope of the smile tends to steepen as the option's life shortens if the adjustment is left undone. As a result, time-adjusted moneyness for each day is calculated as:

$$M_t = 1 + \frac{\frac{K}{S_t} - 1}{\sqrt{T}}$$

where $M_t$ represents time-adjusted moneyness, $K$ is the option strike price, $S_t$ the price of underlying index at time t and $T$ the remaining life of the option. Figures 3 and 4 presents implied volatility curves on 31.1.2018 both before and after time-adjustment for option series' maturing in March and June 2018. Time-to-maturity for the contracts are thus 43 and 143 days, respectively.

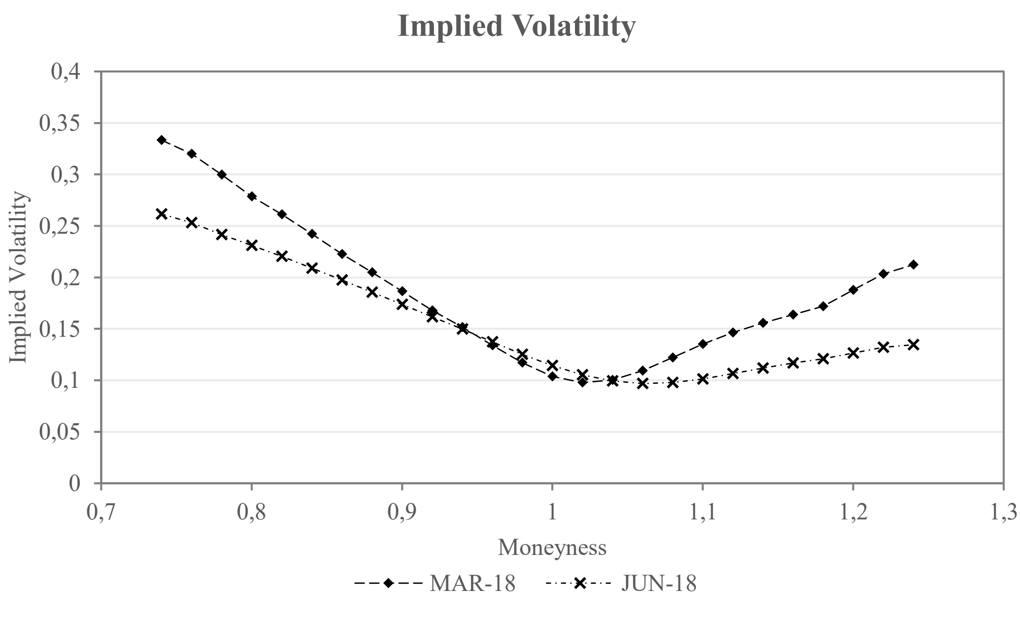



*Figure 3. Implied volatility for March and June 2018 options ranging in moneyness from 0,75 to 1,25*

Consistency in the shape of smiles is visibly improved after the adjustment, as the gap between March and June contracts is greatly narrowed. Furthermore, considerable improvement is seen in the resemblance of slopes as well as the positions of the nadir between the two smiles.

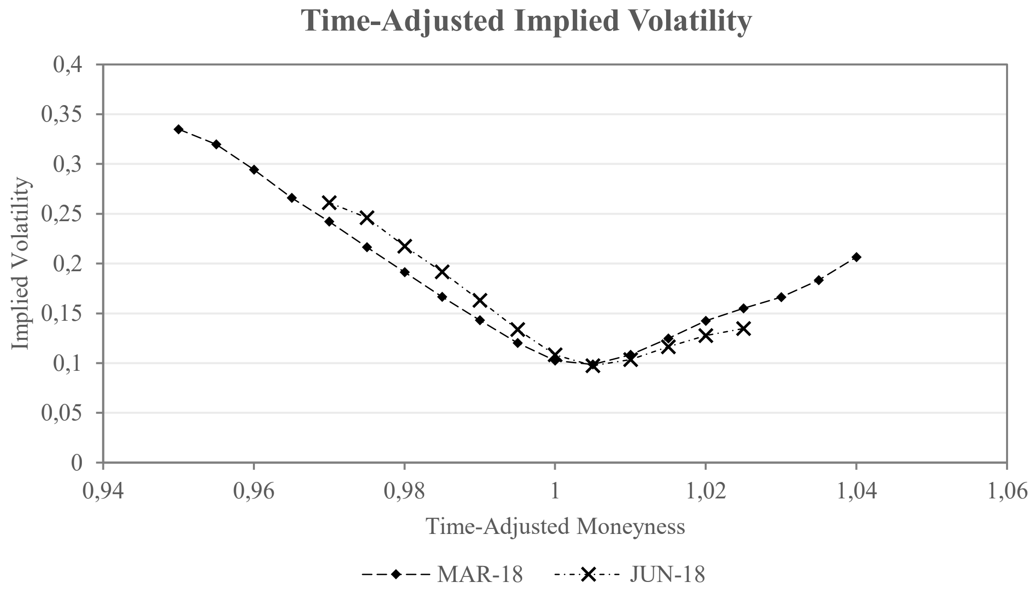



*Figure 4. Time-adjusted smile for the same set of options.*

After daily time-adjusted moneyness for all viable options is calculated, daily implied volatility functions are constructed using a deterministic volatility function, DVF, framework hypothesized for instance by Dupire (1994) and Rubinstein (1994). Similar approach has been adopted by Dumas et al. (1998) and more recently by e.g. Wong and Heaney (2017). Implied volatility functions are derived by regressing IV with time-adjusted moneyness and its square, due to the parabolic nature of the smile. Inclusion of time-to-maturity and its cross-variable with moneyness to the function is considered; however, Dumas et al. (1998) note that this often leads to major overfitting, and thus time-to-maturity is considered only in adjusted moneyness. Daily quadratic implied volatility functions are formulated as:

$$\sigma_{IV,t} = \alpha_0 + \beta_1 M_t + \beta_2 M_t^2 + \omega_t$$

where $\sigma_{IV,t}$ is implied volatility at time t and $M_t$ and $M_t^2$ are option moneyness and its square at time t, calculated in line with Dumas et al (1998). $R^2$ obtained from daily regressions for implied volatility functions are consistently north of 0,9 with the minimum value at 0,932 and average at 0,985. For comparison, volatility functions constructed without time-adjustment with minimum at 0,824 and average at 0,911. First and second derivative of the quadratic volatility function are determined to depict the skewness and convexity of the smile, and thus are used as a proxy for third and fourth moment of the expected price distribution. This is done in line with Wong and Heaney (2017), who build on the work of Bakshi et al. (2003), with similar method adopted by e.g. Buyn and Kim (2013). Risk-neutral higher moments used in this thesis are formulated as:

$$SK_t = \beta_1 + 2M_t\beta_2$$

and

$$KU_t = 2\beta_2$$

where $SK_t$ and $KU_t$ are risk-neutral skewness and risk-neutral kurtosis at time t. As mentioned earlier, derived proxies for risk-neutral skewness and risk-neutral-kurtosis should include information regarding expectations of physical distribution of future index prices (Backus et al., 2004). Charts depicting the development of risk-neutral higher moments are provided in Appendix 1. It is noteworthy that as the time-adjustment of option moneyness impacts the moneyness reported, values of risk-neutral skewness and risk-neutral kurtosis derived from quadratic volatility functions are affected. Before time-adjustment risk-neutral skewness varies between -1,5 and -0,29, and after time-adjustment between -7,0 and -1,95; correlation of sets of risk-neutral skewness is 0,64. Correspondingly, from the original range of -3,5 to 21,7, after time-adjustment values of risk-neutral kurtosis range from -28,6 and 68,0, with the correlation of time-adjusted and normal risk-neutral kurtosis is 0,5.

It is evident that the ability to clearly interpret the actual impact of the variables on future volatility is reduced; however, as noted by Dumas et al. (1998), to ensure the comparability of risk-neutral skewness and kurtosis, moneyness should be adjusted by time. Though time-adjustment done improves the correspondence of the used variables over time, it does not provide a clear-cut solution to the issue related to maturity mismatch problem noted by Lamoureux and Lastrapes (1993): Due to quarterly maturities of S&P 500 options, after option life is less than 10 trading days, option series maturing after the next quarter is used to derive daily quadratic implied functions and variables depicting their shapes, leading to undesirable peaks in the data. In this thesis, the possibly-reduced explanatory power related to these peaks is accepted as an unfortunate feature of S&P 500 options.

Proxies for latent volatility are estimated with Yang-Zhang estimator, introduced in chapter 2.3. YZ is preferred as a volatility proxy since it is a drift-independent estimator of volatility accounting for overnight jumps, thus providing most accurate estimates of volatility in the absence of intraday data (Yang & Zhang, 2000; Shu & Zhang, 2006). Three sets of volatility proxies are estimated using daily, weekly and monthly historical data. This is done (1) to obtain partial volatility components for the HAR-RV model and (2) to enable daily, weekly and monthly volatility forecasts. Daily, weekly and monthly volatilities estimated via Yang-Zhang method are provided in Appendix 2.

In this thesis, the general HAR-RV proposed by Corsi (2009) is used as the benchmark model. General model is then enhanced by including risk-neutral skewness and risk-neutral kurtosis both separately and together. Out-of-sample forecasting performance between general and enhanced models is then evaluated with root mean squared error and mean absolute error to see (1) does the shape of the smile include significant information of future volatility and (2) whether either one of the risk-neutral higher moments dominates the other as a source of this information. The latter question can be interpreted as whether the slope or the convexity of the smile should be preferred as a predictor of future volatility. General model by Corsi (2009) is formulated as:

$$\sigma_{t,T} = \alpha_0 + \beta_1\sigma_{t-1,d} + \beta_2\sigma_{t-5,w} + \beta_3\sigma_{t-21,m} + \omega_t$$

where $\sigma$ denotes realized volatility for future period $T$ forecasted at time $t$, and $d$, $w$ and $m$ represents daily, weekly and monthly periods for which realized volatility is estimated. Models of interest are created in additive manner, and will be referenced as HAR-SK for risk-neutral skewness, HAR-KU for risk-neutral kurtosis, and Moment-HAR, including both third and fourth risk-neutral higher moments. Enhanced models are formulated as:

$$\sigma_{t,d} = \alpha_0 + \beta_1\sigma_{t-1,d} + \beta_2\sigma_{t-1,w} + \beta_3\sigma_{t-1,m} + \theta_1 SK_{t-1} + \omega_t$$

$$\sigma_{t,w} = \alpha_0 + \beta_1\sigma_{t-1,d} + \beta_2\sigma_{t-1,w} + \beta_3\sigma_{t-1,m} + \gamma_1 KU_{t-1} + \omega_t$$

$$\sigma_{t,m} = \alpha_0 + \beta_1\sigma_{t-1,d} + \beta_2\sigma_{t-1,w} + \beta_3\sigma_{t-1,m} + \theta_1 SK_{t-1} + \gamma_1 KU_{t-1} + \omega_t$$

Additional benchmark model, HAR-RV-IV, is also used, where implied volatility is added as an regressor in a parallel manner. Similar method has been employed by at least Busch, Christensen and Nielsen (2008) and Byun and Kim (2013). This is done to study whether the shape of the smile has additional information compared to plain implied volatility. Implied volatility used in HAR-RV-IV is an average of volatilities implied by options with moneyness from 1,02 to 1,04, as Ederington and Guan (2005) provide evidence that the information content in these options is the greatest.

Correlation matrix of all used variables is provided in Appendix 3. Correlations of risk-neutral skewness with different volatility components, daily, weekly and monthly, are around -0,4, meaning that, in general, increase in risk-neutral skewness, i.e. slope of the smile turns less negative, is accompanied by decrease in volatility. For risk-neutral kurtosis, correlations with volatility components are between -0,65 and -0,7. Thus, as risk-neutral kurtosis increases, i.e. the smile flattens (Bakshi et al., 2003), volatility tends to decrease. Correlation between risk-neutral higher moments is 0,64, implying that changes in the slope and convexity of the smile tend to move somewhat likewise. Correlations between implied volatility and distinct volatility components vary around 0,8. Regarding risk-neutral higher moments, correlation of implied volatility is -0,57 with skewness and -0,81 with kurtosis.

Rolling forecast is selected as a forecasting methodology for this thesis. In-sample data is used to fit initial models, with a period-length of approximately six months, or 126 trading days is chosen to ensure sufficient amount of data for model fitting. With the fitted models, daily, weekly and monthly volatility predictions are made on a daily basis. For instance, regarding monthly predictions, dependent variable is the monthly annualized volatility estimated with Yang-Zhang 22 days ahead, representing monthly volatility for the next month. After volatility prediction for certain horizon is done, the forecasting horizon is moved ahead by one day. Rolling forecasts are accompanied by weekly recalibrations of the model as suggested by Brownlees, Engle and Kelly (2011), to account for parameter drift. For model recalibration, the length of the estimation period is equal to the length of in-sample data, that is, models are re-estimated with 126 most recent observations. Forecasting procedure is continued until the end of out-of-sample data is reached, after which predictions are evaluated against realized volatility during distinct forecast horizons.

# 5. Empirical Results

This chapter provides an overview and in-depth analysis of the results of the study. Performance of daily, weekly and monthly volatility forecasts with baseline and modified HAR-RV-models are presented together with statistics regarding model significance with confidence level of 95%. In addition to analysis during the whole period, forecasts during and excluding the crash period with high volatility from 21.2.2020 to 14.4.2020 is presented to showcase model performance during dissimilar market periods. Daily forecasts are analysed first, followed by corresponding analyses for weekly and monthly forecasts, after which empirical results are summarized and compared to earlier literature.

## 5.1. Daily forecasts

Table 2 provides an overview of daily forecasting performance with enhanced HAR-RV models. Of the used models, HAR-IV, the benchmark model enhanced with implied volatility, has performed the best during the whole period according to total RMSE loss function value. When comparing mean of absolute errors, baseline HAR-RV model has provided most accurate forecasts, albeit only by slight margin. Realized daily volatility, forecasts with HAR-IV and forecasting errors are shown in Figure 5; similar charts for other models are provided in Appendices 4 to 9. The second most accurate model by RMSE is HAR-IV-KU, an HAR-IV model enhanced with risk-neutral kurtosis, followed by Moment-HAR model including both risk-neutral higher moments as additional regressors. Justification for enhancing HAR-IV model with risk-neutral kurtosis but not skewness is provided shortly. Regarding daily forecasts, according to RMSE, the worst performance during the period is coming from the model with risk-neutral skewness as an additional regressor, HAR-SK; however, regarding MAE values, it overperforms other model enhanced with risk-neutral higher moments, apart from HAR-IV-KU.

*Table 2: Daily forecasting performance*

**Forecasting performance, daily**

| Model | Whole period | | Excluding crash | | During crash | | St. Dev of forecast error |
|---|---|---|---|---|---|---|---|
| | **RMSE** | **MAE** | **RMSE** | **MAE** | **RMSE** | **MAE** | |
| *HAR-RV* | *7,38* | ***3,95*** | *4,56* | ***3,18*** | *28,26* | *20,81* | *8,15* |
| *HAR-SK* | *7,54* | *4,04* | *4,61* | *3,23* | *29,05* | *21,86* | *8,33* |
| *HAR-KU* | *7,33* | *4,13* | *4,61* | *3,32* | *27,75* | *21,97* | *8,08* |
| *Moment-HAR* | *7,09* | *4,07* | *5,02* | *3,37* | ***24,61*** | ***19,56*** | *7,82* |
| *HAR-IV* | ***6,79*** | *3,98* | ***4,54*** | *3,26* | *24,70* | *19,81* | *7,40* |
| *HAR-IV-KU* | *7,04* | *4,04* | *4,60* | *3,27* | *25,99* | *20,99* | *7,70* |

When excluding the crash period in the first quarter of 2020, both benchmark models overperform against the models enhanced with the shape of the volatility smile, with HAR-IV having the lowest RMSE and HAR-RV the lowest MAE. According to RMSE values, when excluding the crash, there seems to be no difference in the forecasting accuracy of models where either risk-neutral skewness or risk-neutral kurtosis is used as an additional regressor. When comparing absolute errors, HAR-SK overperforms against other smile models. Moment-HAR, the model including both risk-neutral higher moments, performs clearly the worst.

However, during the crash period, Moment-HAR has the lowest loss function values for both RMSE and MAE, though the difference to HAR-IV in RMSE is quite marginal. When analysing the other, less significant period of increased volatility and declining S&P 500 index levels, from 1.10.2018 to 3.1.2019 when the index was down -16,3% and the average daily Yang-Zhang volatility was 20,4%, HAR-IV was the most accurate model with RMSE of 5,70 with Moment-HAR obtaining a value of 6,01. These results imply that regarding daily forecasts, variables depicting the shape of the smile seems to contain some additional explanatory power during periods of plunging index levels and high volatility; however, it should be noted that statistical significance of these results is not tested. This outcome supports results of Doran et al. (2007) and Yan (2011), who find empirical evidence that shape of the smile predicts future stock returns and particularly downward jumps.

Table 3 presents information regarding explanatory power of used models with maximum, minimum mean and median $R^2$ values over all model fittings done weekly during the forecasting period, amounting to 163 recalibrations for daily forecasts. This information is accompanied by coefficient overviews of these recalibrations for all models, which for daily forecasts are presented in Appendices 10 to 15. Baseline HAR-RV has the lowest mean $R^2$ with still a fairly high value of 0,64, accompanied by HAR-SK, implying that for the daily volatility forecasts the risk-neutral skewness provides no additional explanatory power. When analysing coefficients provided in Appendix 10 it is noted that increasing daily and weekly volatilities today have an increasing impact for daily volatility tomorrow, confirming the notion that Heterogenous Autoregressive models are able to capture volatility clustering in the data, as suggested by Corsi (2009). For HAR-RV model, daily volatility today is statistically significant predictor for daily volatility tomorrow 162 out of 163 times when using 95% confidence level; weekly volatility is significant predictor during approximately every third model recalibration.

*Table 3: Overview of explanatory power*

| **$R^2$ over 163 recalibrations, daily** | | | | |
|---|---|---|---|---|
| **Model** | **Maximum** | **Minimum** | **Mean** | **Median** |
| *HAR-RV* | *0,88* | *0,42* | *0,64* | *0,60* |
| *HAR-SK* | *0,88* | *0,43* | *0,64* | *0,60* |
| *HAR-KU* | *0,89* | *0,48* | *0,66* | *0,64* |
| *Moment-HAR* | *0,89* | *0,50* | *0,67* | *0,64* |
| *HAR-IV* | *0,89* | *0,46* | *0,68* | *0,64* |
| *HAR-IV-KU* | *0,90* | *0,48* | *0,69* | *0,67* |

Coefficients in Appendix 12 for HAR-SK model reveal that while risk-neutral skewness seems to have a negative relationship with daily volatility, i.e. increasing skewness and thus the slope of the smile tends to decrease volatility tomorrow, the coefficient is statistically significant for only 11% of model fittings. Hence, the predictive power of risk-neutral skewness regarding daily volatility cannot be confirmed. However, risk-neutral kurtosis seems to have an statistically significant relationship with daily volatility as its coefficient is significant 130 out of 163 recalibrations or 80% of the time. Negative coefficient implies that flattening smile tends to lead to decrease in daily volatility, according to argument by

Bakshi et al. (2003) that higher risk-neutral kurtosis results in a flatter smile. HAR-KU is also better able to explain the variance of daily volatility around its mean, when comparing $R^2$ values to HAR-RV and HAR-SK. Due to this statistically significant relationship, HAR-IV model is also enhanced with risk-neutral kurtosis to see whether the convexity of the smile has incremental predictive power over plain implied volatility.

HAR-IV-KU has the highest mean $R^2$ of all models; however, its forecasting performance is worse when compared to HAR-IV model, implying that, at most, risk-neutral kurtosis has negligent amount of incremental information regarding future daily volatility. Regarding coefficient significance in HAR-IV-KU, implied volatility is significant 71% and risk-neutral kurtosis 42% of recalibrations. HAR-IV has the second highest mean $R^2$ with implied volatility having a positive, statistically significant coefficient 157 out of 163 recalibrations. Mean $R^2$ for Moment-HAR is the third highest with 0,67, a result in line with forecasting accuracy in Table 2. Regarding coefficients, risk-neutral skewness is statistically significant 23% and risk-neutral kurtosis 74% of recalibrations.

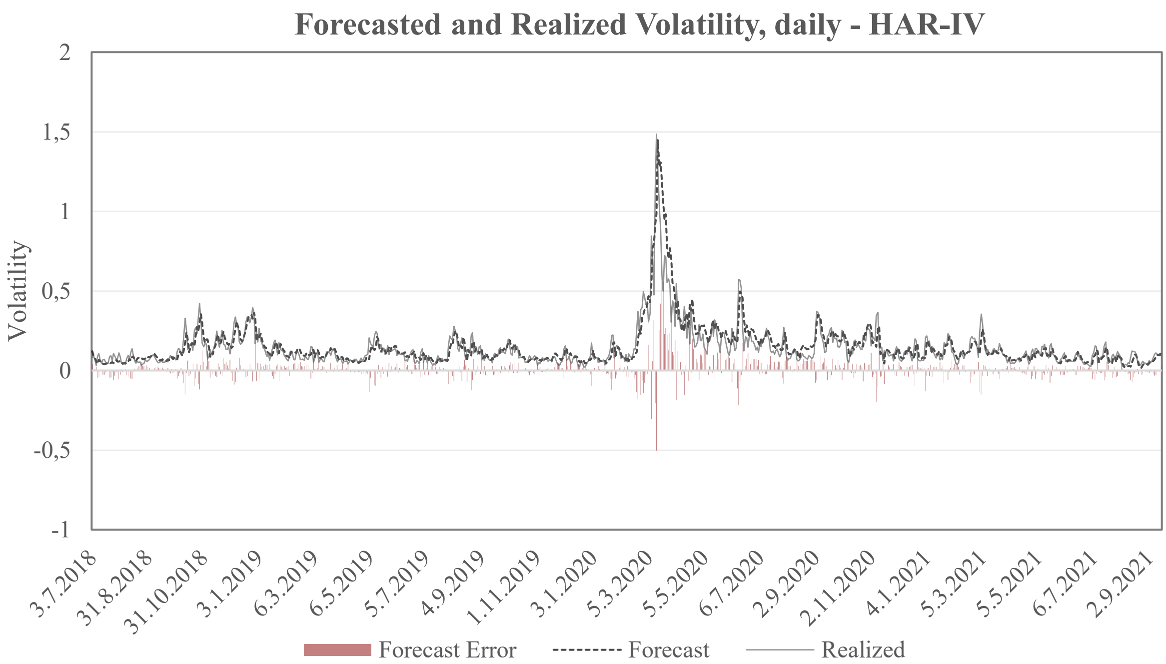



*Figure 5: Realized and forecasted daily volatility with HAR-IV*

## 5.2. Weekly forecasts

Results of weekly volatility forecasts are presented in Table 4. For the whole period, forecasts with Moment-HAR model are clearly the most accurate according to RMSE loss function value; however, its predictions are second best after HAR-IV benchmark when comparing absolute errors. Forecasted and actual volatilities for Moment-HAR are portrayed in Figure 6; for other models, these are provided in Appendices 4 to 9. Of the models enhanced with a single risk-neutral higher moment, HAR-KU with second-best forecasting accuracy according to RMSE dominates HAR-SK, which has the worst performance of all models as seen in the highest values for both RMSE and MAE. Including risk-neutral kurtosis as an additional regressor in the HAR-IV model increases the forecasting performance according to RMSE when compared to baseline implied volatility model, a feature not seen in the daily volatility forecasts. However, this edge is lost in MAE.

*Table 4: Weekly forecasting performance*

**Forecasting performance, weekly**

| Model | Whole period | | Excluding crash | | During crash | | St. Dev of forecast error |
|---|---|---|---|---|---|---|---|
| | **RMSE** | **MAE** | **RMSE** | **MAE** | **RMSE** | **MAE** | |
| *HAR-RV* | *9,32* | *4,78* | *5,31* | *3,92* | *37,09* | *23,68* | *10,28* |
| *HAR-SK* | *9,48* | *4,96* | *5,48* | *4,04* | *37,44* | *25,33* | *10,40* |
| *HAR-KU* | *8,07* | *4,79* | *5,41* | *3,95* | *29,22* | *23,25* | *8,79* |
| *Moment-HAR* | ***7,67*** | *4,67* | *5,60* | *3,95* | ***25,74*** | ***20,44*** | *8,35* |
| *HAR-IV* | *9,00* | ***4,59*** | ***5,17*** | ***3,76*** | *35,69* | *22,83* | *9,86* |
| *HAR-IV-KU* | *8,81* | *4,81* | *5,42* | *3,84* | *33,70* | *25,94* | *9,65* |

When excluding the crash period of 2020, benchmark models HAR-IV and HAR-RV have performed the best, providing evidence of the redundant nature of risk-neutral higher moments as explanatory variables during more or less stable markets. This notion is strengthened by the worst forecasting accuracy with Moment-HAR according to RMSE. Of the models enhanced with risk-neutral higher moments, HAR-KU performs the best according to RMSE, with HAR-IV-KU having the lowest MAE of these models.

However, the importance of risk-neutral higher moments, particularly kurtosis, is seen in the forecasting accuracy during the crash period. Moment-HAR provides undoubtedly the most accurate volatility forecasts in terms of both RMSE and MAE, followed by HAR-KU. Worst forecasting performance during the crash comes from HAR-SK, providing some evidence of the superiority of risk-neutral kurtosis over risk-neutral skewness. Regarding the milder high volatility period during the last quarter of 2018, HAR-IV provides the most accurate forecasts with RMSE of 7,05, followed by Moment-HAR and HAR-IV-KU, having RMSE values of 7,18. RMSE value of HAR-KU is 7,20, with HAR-SK and baseline HAR-RV performing worst with RMSE values of 7,57 and 7,62, respectively. These results provide additional evidence regarding the importance of risk-neutral kurtosis during a jumpy period in the market, already seen in daily forecasts and confirming the results of Doran et al. (2007) and Yan (2011).

*Table 5: Overview of explanatory power*

| **$R^2$ over 162 recalibrations, weekly** | | | | |
|---|---|---|---|---|
| **Model** | **Maximum** | **Minimum** | **Mean** | **Median** |
| *HAR-RV* | *0,86* | *0,19* | *0,49* | *0,48* |
| *HAR-SK* | *0,87* | *0,23* | *0,50* | *0,49* |
| *HAR-KU* | *0,92* | *0,30* | *0,56* | *0,52* |
| *Moment-HAR* | *0,92* | *0,31* | *0,57* | *0,52* |
| *HAR-IV* | *0,89* | *0,29* | *0,57* | *0,56* |
| *HAR-IV-KU* | *0,92* | *0,31* | *0,60* | *0,57* |

Table 5 presents again information regarding distributions of $R^2$ of weekly model recalibrations, amounting to a total of 162 model fittings. Mean $R^2$ values range from 0,49 of baseline HAR-RV to 0,6 of HAR-IV-KU. Additional information regarding coefficient significance in fitted models are provided in Appendices 10 to 15. HAR-RV has the worst model fit again, and the negligible improvement in $R^2$ for HAR-SK showcases the lack of importance of risk-neutral skewness in improving the model fit, as was the case with daily forecasts. HAR-IV-KU has the highest explanatory power, followed by Moment-HAR, HAR-IV and HAR-KU with comparable model fits.

Impact of different volatility components can be seen when examining coefficient overviews presented in Appendices 10 to 15. Regarding HAR-RV, both daily and weekly volatility components again have positive mean coefficients, implying the ability of the model to account for volatility clustering. However, the component for monthly volatility has a negative mean coefficient, albeit only slightly. Over the course of 162 weekly model recalibrations, daily component is statistically significant 161 times (99%), weekly component 79 times (48%) and monthly component 63 times (39%), portraying the increased importance of longer-term volatility components regarding weekly volatility forecasts. Regarding models enhanced with implied volatility, coefficient for IV is statistically significant 135 times out of 162, or 83%, in HAR-IV, and 89 times, or 55%, in HAR-IV-KU. Coefficient for risk-neutral kurtosis in this model is significant 86 times, amounting to 53% of total number of model recalibrations.

Regarding Moment-HAR model, risk-neutral kurtosis once again dominates risk-neutral skewness in the significance during model fitting. Risk-neutral kurtosis is statistically significant for 124 or 76% of model fittings, with risk-neutral skewness having significance 56 times, i.e. about every third model recalibration. Coefficients for risk-neutral higher moments are significant 129 times out of 162 for HAR-KU, or 79%, and 64 time for HAR-SK, amounting to 41% of total number of recalibrations. Coefficients of risk-neutral kurtosis are again negative, which can be interpreted as increase in kurtosis, i.e. flattening smile (Bakshi et al., 2003), leads to a decline in weekly volatility. For risk-neutral skewness the mean of coefficients is positive, implying that volatility decreases when the slope of the smile steepens, a result in contradiction with earlier literature. It is clear that the statistical relationship with weekly volatility is clearly more prominent for risk-neutral kurtosis than for risk-neutral skewness, as was the case with daily volatility. This finding coupled with the superior forecasting accuracy of HAR-KU against HAR-SK provides evidence of the dominance of risk-neutral kurtosis over skewness as an additional predictor for weekly volatility. This implies that the convexity rather than the slope of the volatility smile contains more information of future volatility.

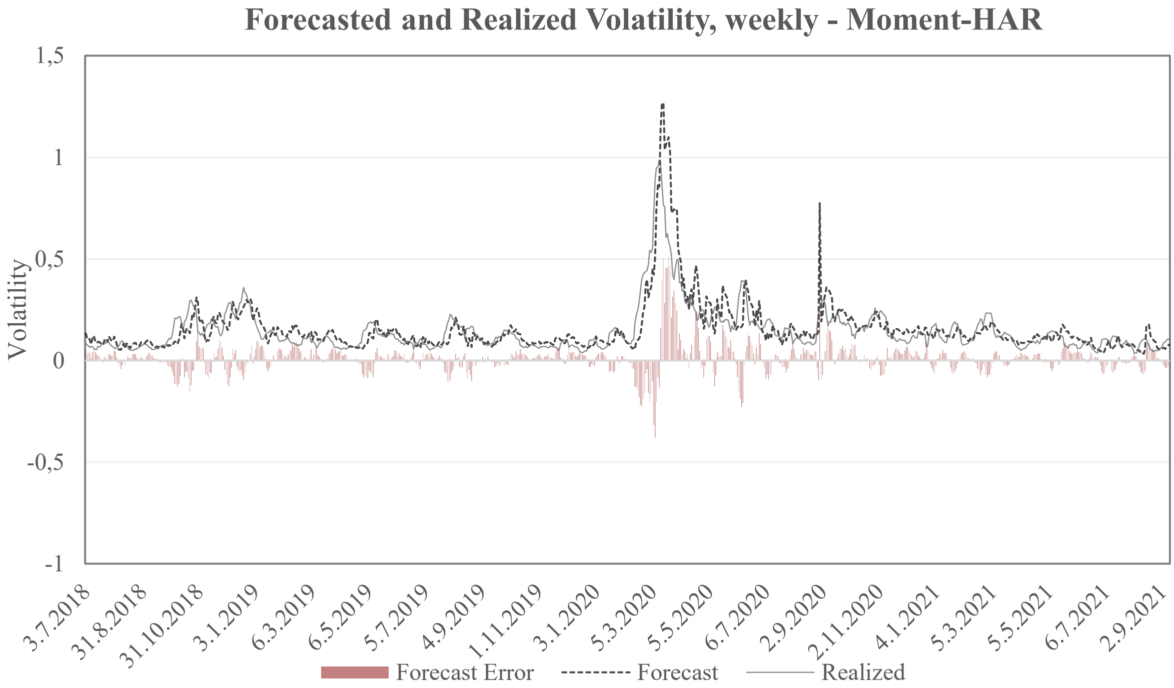



*Figure 6: Weekly volatility forecasts with Moment-HAR*

## 5.3. Monthly forecasts

Finally regarding the monthly volatility forecasts, the details of forecasting accuracy are provided in Table 6. According to values of RMSE and MAE loss functions, HAR-IV-KU overperforms other models in forecasting accuracy during the whole period. Forecasted and actual volatilities for HAR-IV-KU are portrayed in Figure 7; for other models, these are provided in Appendices 4 to 9. According to RMSE, second- and third-best performances come from Moment-HAR and HAR-KU models, reaffirming the role of risk-neutral kurtosis in volatility forecasting. Interestingly, HAR-IV, which has been one of the top performers in daily and volatility forecasting, has the poorest performance in monthly forecasts over the whole period in terms of RMSE. Additionally, HAR-SK model underperforms against the baseline HAR-RV model, providing further evidence of its redundancy in volatility forecasting at least during this period, a result similar to those of daily and weekly volatility forecasts.

*Table 6: Monthly forecasting performance*

**Forecasting performance, monthly**

| **Model** | Whole period | | Excluding crash | | During crash | | **St. Dev of forecast error** |
|---|---|---|---|---|---|---|---|
| | **RMSE** | **MAE** | **RMSE** | **MAE** | **RMSE** | **MAE** | |
| *HAR-RV* | *12,57* | *6,19* | *8,17* | *5,20* | *46,04* | ***26,41*** | *13,82* |
| *HAR-SK* | *12,86* | *6,47* | *8,34* | *5,34* | *47,20* | *30,74* | *14,09* |
| *HAR-KU* | *11,73* | *6,28* | *8,25* | *5,19* | *40,38* | *29,66* | *12,77* |
| *Moment-HAR* | *11,43* | *6,05* | *8,16* | *5,02* | ***38,83*** | *28,31* | *12,50* |
| *HAR-IV* | *13,00* | *5,87* | *7,99* | *4,80* | *49,31* | *28,91* | *14,39* |
| *HAR-IV-KU* | ***11,32*** | ***5,72*** | ***7,95*** | ***4,73*** | *39,03* | *26,84* | *12,56* |

Excluding the crash period, HAR-IV-KU is still the top performer, albeit only by a small margin to the forecasting accuracy of HAR-IV. Of the models enhanced with risk-neutral higher moments, only Moment-HAR tops the baseline HAR-RV, though only by a negligible difference in RMSE and by slightly larger difference in MAE, and underperforms against HAR-IV benchmark. However, for the monthly volatility forecasts, the shape of the volatility smile seems to contain at least some incremental information regarding future monthly volatility even during the neutral period excluding the crash.

During the crash, models overperforming are again the ones including risk-neutral kurtosis, with Moment-HAR and HAR-IV-KU providing the most accurate forecasts according to RMSE and MAE, followed by and HAR-KU. Worst performance by RMSE values is shown by plain implied volatility model HAR-IV, followed by HAR-SK and HAR-RV. When analysing the less prominent high-volatility period during last quarter of 2018, HAR-IV-KU is again the best-performing model, with RMSE loss function value of 6,94, providing evidence of the importance of risk-neutral kurtosis in monthly volatility forecasting during adverse markets; RMSE value for HAR-IV is 7,82 for the period. Furthermore, the second-best performance during this subperiod comes from HAR-KU with RMSE of 7,37, followed by Moment-HAR with RMSE of 7,38. These results further confirm the notion by Doran et al. (2007) and Yan (2011) that the shape of the volatility smile seems to have incremental predictive power regarding future volatility, at least during adverse periods.

*Table 7: Overview of explanatory power*

| **$R^2$ over 157 recalibrations, monthly** | | | | |
|---|---|---|---|---|
| **Model** | **Maximum** | **Minimum** | **Mean** | **Median** |
| *HAR-RV* | *0,71* | *0,03* | *0,26* | *0,27* |
| *HAR-SK* | *0,73* | *0,03* | *0,31* | *0,32* |
| *HAR-KU* | *0,78* | *0,05* | *0,36* | *0,36* |
| *Moment-HAR* | *0,80* | *0,05* | *0,39* | *0,39* |
| *HAR-IV* | *0,78* | *0,08* | *0,41* | *0,40* |
| *HAR-IV-KU* | *0,79* | *0,09* | *0,47* | *0,48* |

Table 7 presents the information of model fits over weekly model recalibrations, amounting to 157 model fittings with monthly forecasting. In addition to forecast accuracy, superiority of HAR-IV-KU regarding monthly volatility is evident in the mean $R^2$ values of used models, with the highest mean $R^2$ of 0,47, followed by 0,41 of plain-implied volatility model HAR-IV. Overviews of coefficients and their significance is provided in Appendices 10 to 15. For HAR-IV-KU, implied volatility is statistically significant 106 out of 157 model fittings, or 65%, and risk-neutral kurtosis 91 times, or 56%. Regarding the baseline HAR-RV model, monthly volatility component has now the highest amount of significant coefficients with 79 (48%). Daily and weekly volatility components have significant components 74 (45%) and 38 (23%) times out of 157 model fittings, respectively. Thus, monthly volatility component seems to be the most significant component in HAR-RV model, albeit only with a slight difference to the daily component.

Regarding risk-neutral higher moments, risk-neutral kurtosis again tends to dominate skewness both in forecasting accuracy and model fit. When fitting Moment-HAR, the coefficient for risk-neutral kurtosis is statistically significant 104 times (64%), compared to 63 times (39%) for risk-neutral skewness. Signs of the coefficient are parallel to those of the corresponding models for daily and weekly volatility, that is, the mean of coefficients is negative for risk-neutral kurtosis and positive for risk-neutral skewness. Again, by interpreting in line with Bakshi et al. (2003), flattening of the smile tends to result in declining monthly volatility, and the steepening slope of the smile tends to be followed by decreasing future monthly volatility. Results regarding forecasting accuracy and statistical significance of coefficients imply that risk-neutral kurtosis is preferred over risk-neutral skewness in monthly volatility forecasting, a result similar to those obtained from daily and monthly forecasts.

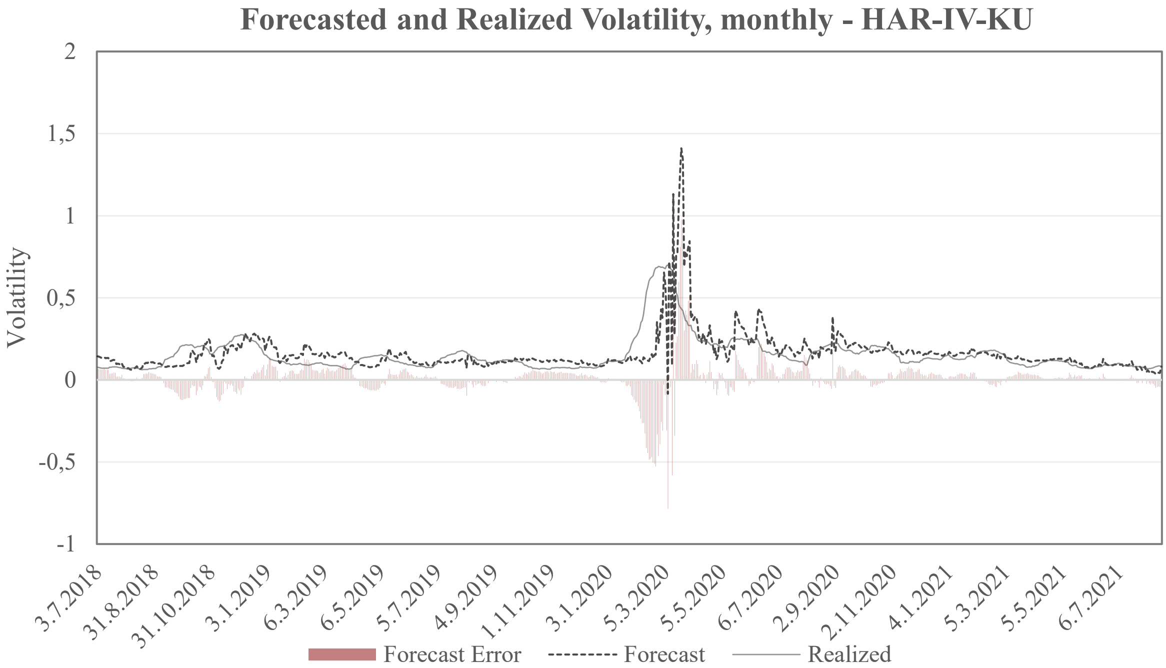



*Figure 7: Realized and forecasted monthly volatility with HAR-IV-KU*

## 5.4. Summary

Results presented above provide evidence of the incremental predictive information contained in the shape of the volatility smile especially during adverse periods, with the shape depicted by risk-neutral skewness and risk-neutral kurtosis derived from the quadratic implied volatility functions. Similar approach to enhancing volatility models has been adopted by at least Byun and Kim (2013) and Wong and Heaney (2017), with the method built on the framework provided either by Bakshi et al. (2003) and Backus et al. (2004). In addition to volatility forecasts, relationship between the shape of the smile and future returns has been studied by at least Doran et al. (2007), Ying, Zhang and Zhao (2010) and Yan (2011), who all provide evidence that highly skewed implied volatility smile tends to be followed by lower future returns. Doran et al. (2007) and Yan (2011) suggest that shape of the smile seems to contain at least some predictive power of future spikes and crashes in the stock prices, and the slope of the skew is used as a proxy for average jump size by Yan (2011). Pena et al. (1999) provide evidence that more convex smile tends to be associated with periods of higher volatility.

The benchmark model enhanced with plain implied volatility produces lowest forecasting errors when predicting daily volatility, suggesting that the accuracy of realized volatility models, at least the HAR-RV by Corsi (2003; 2009), can be improved by inclusion of option implied volatility. HAR-IV overperforms other models during the period as well as when the crash of 2020 is excluded, and also provides the best accuracy over a three-month period of high volatility and plunging index levels during the last quarter of 2018. It is also the second-most accurate model during the crash of 2020 after the Moment-HAR, a model enhanced by both risk-neutral higher moments, providing robust evidence of its superiority in daily volatility forecasting. Similar results have been obtained by Busch et al. (2008) and Byun and Kim (2013), who find evidence that including implied volatility improves the performance of HAR-RV for daily, weekly and monthly forecasting horizons. Relatively good accuracy of Moment-HAR during adverse periods in markets provide evidence complementing the work of Doran et al. (2007) and Yan (2011), who suggest that the shape of the volatility smile contains some predictive information of future spikes and crashes. However, the model performs worst when excluding the crash period, suggesting that risk-neutral higher moments are redundant for models forecasting daily volatility.

Regarding model fit and significance, inclusion of risk-neutral kurtosis in the implied volatility model improves the explanatory power only slightly for daily forecasts, increasing the mean $R^2$ of all fitted models by 0,01, and decreases the significance of implied volatility in the model. Thus, according to Byun and Kim (2013), it could be concluded that implied volatility already includes most of the information of risk-neutral kurtosis. These results suggest that between 2018 and 2021, risk-neutral higher moments of the underlying future price distribution contain at most a negligible amount of information regarding future volatility, a result contradicting those of Byun and Kim (2013).

Evidence on importance of risk-neutral higher moments, particularly risk-neutral kurtosis, is obtained from results of weekly volatility forecasts. During the full period, Moment-HAR provides the most accurate predictions, followed by HAR-KU, a model including only risk-neutral kurtosis as an additional regressor. Same models provide the most accurate predictions during the crash period, further confirming the results of Doran et al. (2007) and

Yan (2011) regarding the predictive information in the smile concerning future jumps. Excluding the crash, the best performance is obtained by benchmark models HAR-IV and HAR-RV, a corresponding result to that obtained from daily forecasts. Of the enhanced models, HAR-KU is the most accurate model when the crash period is excluded.

Highest mean explanatory power is obtained by HAR-IV-KU model with R2 at 0,6, followed by Moment-HAR, HAR-IV and HAR-KU with 0,57, 0,57 and 0,56, respectively. When comparing the significance of coefficients of risk-neutral higher moments, it is evident that risk-neutral kurtosis dominates risk-neutral skewness in statistical significance, and when including both components, the significance of risk-neutral kurtosis is decreased, which again could be interpreted that risk-neutral kurtosis includes most if not all information regarding future weekly volatility contained in risk-neutral skewness (Byun & Kim, 2013). In addition to model fit and statistical significance, HAR-KU dominates HAR-SK in forecasting accuracy, implying that it is the convexity rather than the slope of the smile that should be used in weekly volatility forecasts, a result contradicting Byun and Kim (2013), whose results suggest vice versa. In addition, some remarks can be made when analysing the signs of the coefficients. The coefficient for risk-neutral kurtosis is negative, suggesting that increase in kurtosis tends to be followed by a declining weekly volatility. Bakshi et al. (2003) show that in index option volatility smiles, where negative skewness is present, increasing kurtosis leads to a more flatter, less convex smile. Thus, it follows that flattening smile in index options tends to be associated with declining volatility the following week, a result contradicting Pena et al. (1999), who conclude that flatter smile is more commonly associated with high-volatility periods.

For monthly volatility forecasts, HAR-IV-KU, the model employing both plain implied volatility and risk-neutral kurtosis, is preferred for the full period as well as when excluding the crash of 2020. Second-best forecasting accuracy is obtained by Moment-HAR for the full period and by HAR-IV when the crash is excluded. During adverse periods, both the crash of 2020 and the less prominent period in 2018, the best performance is coming from HAR-IV-KU and Moment-HAR, providing evidence that the shape of the smile provides incremental information of sudden increases in future volatility also with monthly forecasting horizon, confirming the results of Doran et al. (2007) and Yan (2011). When comparing only the models enhanced with a single risk-neutral higher moment, HAR-SK and HAR-KU, forecasting accuracy of HAR-KU dominates that of its counterpart during the full period, adverse markets and when excluding the crash period. Superiority of risk-neutral kurtosis over risk-neutral skewness in predicting monthly volatility contradicts the results of Byun and Kim (2013). Furthermore, Wong and Heaney (2017) find no evidence of dominance of either risk-neutral higher moment with monthly forecast period, a result that cannot be confirmed by this thesis.

In addition to its forecasting accuracy with one-month horizon, HAR-IV-KU obtains the greatest explanatory power by some margin with mean R2 of 0,47, compared to 0,41 of HAR-IV and 0,39 of Moment-HAR. The best model fit coupled with the most accurate forecasting performance over the full period and different subperiods provide evidence that models forecasting monthly volatility should be enhanced by including risk-neutral higher kurtosis and implied volatility into the model. Coefficient signs imply that flattening of the volatility smile tends to be followed by decreasing monthly volatility, as was the case with weekly volatility. Furthermore, increase in implied volatility tends to be followed by increasing realized volatility, in line with earlier literature (Christensen & Prabhala, 1998; Blair et al., 2001).

Hence, it seems that at least risk-neutral kurtosis has contained predictive information about future volatility between 2018 and 2021 for weekly and monthly forecasting horizons, in some cases even exceeding the information of option implied volatility. Implied volatility used in the models is an average of all options ranging in moneyness between 1,02 and 1,04, as Ederington and Guan (2005) suggest these options have the highest information content regarding future volatility. Results from weekly and monthly forecasts provides confirmation of the relationship demonstrated by Bakshi et al. (2003) and Backus et al. (2004), with changes in the expected density of future prices of the underlying are seen as the changes of the shape of the implied volatility smile. This thesis provides evidence that these the shape of the smile can be used to improve the accuracy of weekly and monthly volatility forecasts, complementing results in earlier literature, for example in Yan (2011), Byun and Kim (2013) and Wong and Heaney (2017). It is noteworthy though to distinguish between results of this thesis and those of Byun and Kim (2013), whose results favour risk-neutral skewness over kurtosis. As noted by Ederington and Guan (2005), for the net buying pressure hypothesis by Bollen and Whaley (2004) to hold, that is, the shape of the smile is driven hedging pressure in non-at-the-money options, there does not need to be any predictive power in the shape of the smile. Even though results in this thesis advocate against this, the said results are somewhat contradictory, particularly during distinct market periods, no comment is made regarding the credibility of net buying pressure hypothesis.

# 6. Conclusions

Often thought as a common measure of risk inherent in the asset, volatility and its forecasting plays a central role in financial domain, including fields such as risk management, portfolio theory and option pricing. Rather than being normally distributed, financial returns and volatility tend to exhibit some commonly observed features, such as leptokurtic, i.e., fat-tailed distribution as well as volatility clustering, with observations of high volatility having a tendency to be followed by subsequent observations of high volatility, often thought to result a leptokurtic distribution. Thus, significant amount of research has been done in the past to formulate models that are able to replicate common features in financial data. One such model is the Heterogenous Autoregressive model for Realized Volatility by Corsi (2003; 2009), who has built the model on the basis of heterogenous market hypothesis by Müller (1993). HAR-RV models volatility with partial components of past volatility over periods differing in length, namely daily, weekly and monthly. As such, the model is parsimonious in nature and can be extended with additional regressors in a straightforward manner. Furthermore, despite its simplicity and lack of proper long-memory features, it is able to replicate commonly observed features in financial data and acts like a true long-memory model. Due to this, the model has been employed and extended for instance by Andersen et al. (2007), McAleer and Medeiros (2008) and Byun and Kim (2013).

Ground-breaking work in option pricing has been provided by Black and Scholes (1973), whose model had an immense effect on option pricing and provided market agents a sophisticated method for pricing options. Black-Scholes employs several variables to price an option, of which the sole unobservable one is the latent volatility of the underlying asset. As such, it has become common practice to infer volatilities from observed option prices, hence called implied volatilities, which are thought to represent the average volatility of the underlying asset over the life of the option. According to Black-Scholes, option implied volatilities for all options with the same underlying and same time to maturity should be equal. However, it is well-observed that options differing only in exercise price still do exhibit deviances in implied volatility, and when plotted against the option moneyness, these implied volatilities often exhibit a smile-like pattern, called the volatility smile. This deviance is often thought to be due to divergence between a lognormal distribution of future underlying prices, as assumed by Black-Scholes, and the physical distribution of the real world, which often tend to exhibit leptokurtosis, as noted above. Previous literature has shown that the volatility smile can be presented in a functional form with quadratic volatility functions, where implied volatility is regressed against option moneyness and its square, as the smile tends to resemble a parabola.

As the volatility smile can be thought to represent market expectations of the future return density during the life of the option, its shape could thus contain some predictive power regarding future returns and volatility. The relationship has been extensively studied, with results by Bakshi et al. (2003), Backus and Foresi (2004) suggesting that the smile and future volatility of an asset in fact tend to exhibit parallel movements with the shape of the smile. Pena et al. (1999) note that higher kurtosis, that is, a more convex smile, is often associated with higher volatility, while Bakshi et al. (2003) find evidence that in the presence of negative skewness, higher kurtosis of the expected underlying distribution tends to result in a flatter smile. As quadratic volatility functions are used to depict daily smiles in this thesis, slopes and convexities, that is, the first and second derivatives, of these functions can be used as proxies for the shape of the underlying return distribution. The method used in this thesis is built on the work of Backus and Foresi (2004) and has been adopted by at least Wong and Heaney (2017).

HAR-RV model by Corsi (2003; 2009) is used in this thesis for reasons mentioned above, and extensions to the model are made to include variables depicting the volatility smile. In addition to HAR-RV, a version extended with average implied volatility of option ranging in moneyness from 1,02 to 1,04 is used as second benchmark model. Choice of the moneyness range is based on the work of Ederington and Guan (2005) concerning the information content in implied volatility and its smile. Models of interest that are extended with variables depicting the smile are HAR-SK with risk-neutral skewness as an additional regressor, HAR-KU which includes risk-neutral kurtosis, and Moment-HAR, including both risk-neutral higher moments. In addition, HAR-IV-KU model is fitted due to the significant relationship between realized volatility and risk-neutral kurtosis, that is, the convexity of the smile, often observed in this thesis. Forecasts of daily, weekly and monthly volatility are performed with the said models, with the obtained results used to answer the research questions of this thesis.

*Does the implied volatility smile contain predictive information regarding future volatility of S&P 500 index for daily, weekly and monthly horizons?*

Results in this thesis imply that at least risk-neutral kurtosis, that is, the convexity of the option implied volatility smile, has contained predictive information about future volatility between 2018 and 2021 for weekly and monthly forecast horizons. According to increased forecasting accuracy in some instances, this information content has exceeded the information of implied volatility of options ranging in maturity from 1,02 to 1,04, a range chosen in line with results of Ederington and Guan (2005) regarding the information content of IV. Overperformance of models extended with variables depicting the volatility smile is recorded in several instances, during distinct periods in the markets as well as with different forecast horizons. It is thus concluded that the smile implied by S&P 500 options has included predictive information regarding future volatility of the index between 2018 and 2021, at least to some extent. This conclusion is elaborated from different perspectives below.

*Does either of risk-neutral third or fourth moments dominate the other as a source of information regarding future volatility?*

To answer the first detailed research question, it is appropriate to compare the forecasting performance and coefficient significance of risk-neutral skewness and risk-neutral kurtosis, representing the slope and the convexity of the smile. Results of daily forecasts are somewhat contradictory. When assessed with RMSE loss function value, HAR-KU overperforms HAR-SK during the whole period as well as during adverse periods in the markets. However, when MAE is used to evaluate the models, HAR-SK provides more accurate forecasts. In addition, both models extended by a risk-neutral higher moment underperforms against the benchmarks, providing evidence of redundancy of these variables in daily volatility forecasting. When assessing the significance of coefficients with a confidence level of 95%, risk-neutral skewness is significant 11% and risk-neutral kurtosis 80% of the model re-estimations. Thus, even though no conclusion can be made of the daily forecasting accuracy, significance of economic relationship with daily volatility clearly favours risk-neutral kurtosis over skewness, a result in contradiction with Byun and Kim (2013), who prefer skewness over kurtosis.

Weekly forecasting results provide evidence of the dominance of risk-neutral kurtosis over risk-neutral skewness, as HAR-KU overperforms HAR-SK when assessed with both RMSE and MAE for the full period, during the crash of 2020, and when excluding the crash. HAR-KU is the second most accurate model after Moment-HAR for the full period. In addition, coefficient significance favours risk-neutral kurtosis again, with significant coefficients seen 76% of re-estimations of Moment-HAR, compared to 31% of risk-neutral skewness. Concluding evidence of the dominance of the convexity over the slope of the smile is seen in the results of monthly volatility forecasts. HAR-KU again overperforms against HAR-SK over all periods and with both loss function values. Furthermore, HAR-IV-KU achieves the most accurate forecasts during all period, advocating for the usage of risk-neutral kurtosis to enhance volatility forecasts. Thus, it is concluded that the convexity of the smile dominates its slope regarding the information content about future volatility, a result contradicting Byun and Kim (2013) who prefer skewness over kurtosis, and Wong and Heaney (2017) who find no evidence of either of the higher moments dominating the other.

*For which future horizon does the information content improve forecasting accuracy the most?*

HAR-IV extension obtains the best accuracy over the whole period and when excluding the crash of 2020, with baseline HAR-RV being the second best when the crash is excluded. Thus, results obtained in this thesis do not advocate for the information content in the smile regarding daily volatility. Results for weekly forecasts provides evidence that it is appropriate to enhance weekly volatility forecasts with variables depicting the shape of the smile, particularly risk-neutral kurtosis representing the convexity of the smile. Best performance in weekly forecasting is achieved by Moment-HAR extension, followed by HAR-KU for the whole period; when excluding the crash, benchmark models have the most accurate forecasts. From the smile-extended models, HAR-KU achieves the highest accuracy when crash of 2020 is extended

Best performance in monthly forecasting was achieved by HAR-IV-KU for the full period as well as when excluding the crash of 2020, followed by Moment-HAR for the full period and by HAR-IV when the crash is excluded. Furthermore, HAR-IV-KU achieves clearly the best model fit with R2 of 0,47, with the second-best fit coming from HAR-IV with R2 of 0,41. These results suggest that for monthly forecasts, the convexity of the smile does indeed contain a considerable amount of information of future volatility, especially when combined with implied volatility itself. Thus, it is concluded that the information content in the smile implied by S&P 500 options has been the most prominent for a monthly horizon between 2018 and 2021.

*Has the information content in the smile improved forecasting accuracy over a period of extreme volatility during Covid-19 crisis?*

Moment-HAR obtains the most accurate performance over the Covid-19 crash both in RMSE and MAE terms, though with only a slight improvement to HAR-IV, which overperforms other models in daily forecasting accuracy during the whole period as well as when excluding the crash of 2020. These results are in line with Doran et al. (2007) and Yan (2011) who find evidence of the information content in the smile regarding future spikes and crashes. However, Moment-HAR has the worst forecasting performance over the whole period of all models considered, including those extended only with either risk-neutral skewness or kurtosis. During the crash of 2020, the best performance in weekly forecasting is obtained by Moment-HAR, followed by HAR-KU, providing evidence that the shape of the smile does indeed contain some predictive content regarding future price jumps.

Additional evidence of the information content in the smile for adverse periods is provided by results for monthly forecasts. Most accurate forecasts are achieved by HAR-IV-KU extension, followed by Moment-HAR, for both the crash of 2020 and the less prominent adverse period during the last quarter of 2018. Performance of HAR-IV-KU combined with the fact that HAR-KU overperforms HAR-SK over the adverse periods provides evidence that it has in fact been the convexity rather than the slope of the smile that contains information of future volatility regarding adverse market periods. Results for all forecasts horizons thus imply that the smile has contained information of future volatility during the period of extreme volatility in 2020.

While it is concluded that the volatility smile has contained some predictive information of the future volatility of S&P 500 index between 2018 and 2021, it is to be noted that this research is subject to several limitations. First, due to the limited availability of option data in Refinitiv Eikon database for S&P 500 options, the period under scrutiny is limited to start only in the beginning of 2018. With the data being available only from the end of 2017 onwards, HAR-RV models had to be fitted with data spanning over half-year periods to be able to forecast volatility during the adverse market period during the last quarter of 2018, a feature highly in the interest of this thesis. While a length of six months for the in-sample dataset is certainly not the shortest, model fit and performance could have been improved with a longer period used in estimation. Second, the issue related to maturity mismatch problem (Lamoureux and Lastrapes, 1993) must be brought forward. With S&P 500 options maturing quarterly, there are periods of significant length where, for instance, one-day-ahead volatility is forecasted using variables depicting the smile of options maturing in two months. While time-adjustment for option moneyness is done in line with Dumas et al. (1998), this does not resolve all issues related to maturity mismatch. Third, in relation to above, variables for depicting the smile are derived for only daily periods, meaning that no smoothing over a longer period has been done, thus contributing to possible inconsistencies in the data. Fourth, statistical significance of differences in forecasting accuracy is left untested, with significance only considered regarding model coefficients. Finally, neither asymmetries in volatility nor the autocorrelation of residuals has been considered; however, this is limitation relates more to the overall fit of the model and not to the smile per se.

Results in this thesis suggest that volatility forecasting models can be improved by including the variables depicting the volatility smile, at least to some extent. Framework and methodology in this thesis provide multiple interesting suggestions for further research. To assess whether results obtained in this study hold across markets, similar studies could be conducted for different stock indices and for individual stocks, or in exchange rate or interest rate markets. However, liquidity of the option market in question should be evaluated to ensure the validity of results. To tackle the maturity mismatch problem, additional research is worthy to consider in markets where options are maturing more frequently. Related to this, methodology for smoothing the daily changes in the shape of the volatility smile could be employed to try to tackle the issue of periodical risk-neutral higher moments. Finally, to account for the volatility of volatility, variance functions such as GARCH could be fitted to HAR-RV residuals, to improve the model fit. This consideration, and others above, are left for further research.

# References

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-Based Estimation of Stochastic Volatility Models. *Journal of Finance*, 57, 1047–1091.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The review of economics and statistics*, 89(4), 701-720.

Audrino, F., & Hu, Y. (2016). Volatility forecasting: Downside risk, jumps and leverage effect. *Econometrics*, 4(1), 8.

Backus, D. K., Foresi, S., & Wu, L. (2004). Accounting for biases in Black-Scholes. Available at SSRN 585623.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of finance*, 52(5), 2003-2049.

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1), 101-143.

Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking & Finance*, 5(3), 363-381.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3), 637-654.

Black, F. (1976). Studies of Stock Price Volatility Changes. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section, 177–181.

Blair, B. J., Poon, S. H., & Taylor, S. J. (2001). Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. In *Handbook of quantitative finance and risk management* (pp. 1333-1344). Springer, Boston, MA.

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

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307-327.

Brooks, C. (2005). Introductory Econometrics for Finance. Cambridge, Cambridge University Press.

Brownlees, C., Engle, R. and Kelly, B. (2011). A Practical Guide to Volatility Forecasting through Calm and Storm. *The Journal of Risk*, 14(2), pp. 3–22.

Busch, T.; Christensen, B. J.; Nielsen, M. Ø. (2008) : The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets, Queen's Economics Department Working Paper, No. 1181, Queen's University, Department of Economics, Kingston (Ontario)

Byun, S. J., & Kim, J. S. (2013). The information content of risk-neutral skewness for volatility forecasting. *Journal of Empirical Finance*, 23, 142-161.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *The Review of Financial Studies*, 6(3), 659-681

Chernov, M., Gallant, A. R., Ghysels, E., & Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2), 225-257.

Chiras, D. P., & Manaster, S. (1978). The information content of option prices and a test of market efficiency. *Journal of Financial Economics*, 6(2-3), 213-234.

Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of financial economics*, 50(2), 125-150.

Christoffersen, P., Elkamhi, R., Feunou, B., & Jacobs, K. (2010). Option valuation with conditional heteroskedasticity and nonnormality. *The Review of Financial Studies*, 23(5), 2139-2183.

Corrado, C. J., & Su, T. (1997). Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *The European Journal of Finance*, 3(1), 73-85.

Corsi, F. (2003). A Simple Long Memory Model of Realized Volatility. Manuscript, University of Southern Switzerland

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2), 174-196.

Corsi, F., Pirino, D., & Reno, R. (2008). Volatility forecasting: The jumps do matter. Università di Siena.

Day, T. E., & Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics*, 52(1-2), 267-287.

Dennis, P., & Mayhew, S. (2002). Risk-neutral skewness: Evidence from stock options. *Journal of Financial and Quantitative Analysis*, 37(3), 471-493.

Derman, E., & Kani, I. (1994). The volatility smile and its implied tree. *Goldman Sachs Quantitative Strategies Research Notes*, 2, 45-60.

Di Matteo, T. (2007). Multi-scaling in finance. *Quantitative finance*, 7(1), 21-36.

Doran, J. S., Peterson, D. R., & Tarrant, B. C. (2007). Is there information in the volatility skew?. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 27(10), 921-959.

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

Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1), 18-20.

Ederington, L., & Guan, W. (2005). The information frown in option prices. *Journal of Banking & Finance*, 29(6), 1429-1457.

Garman, M., and Klass, M. (1980). On the Estimation of Security Price Volatilities from Historical Data. *Journal of Business*, 53, 67–78.

Giot, P. (2005). Implied volatility indexes and daily value at risk models. *The Journal of derivatives*, 12(4), 54-64.

Gutierrez, J. A., Martinez, V., & Tse, Y. (2009). Where does return and volatility come from? The case of Asian ETFs. *International Review of Economics & Finance*, 18(4), 671-679.

Hull, J. C. 2017. Options, Futures and Other Derivatives. 10th ed. Upper Saddle River: Prentice Hall.

Kambouroudis, D. S., McMillan, D. G., & Tsakou, K. (2016). Forecasting stock return volatility: A comparison of GARCH, implied volatility, and realized volatility models. *Journal of Futures Markets*, 36(12), 1127-1163.

Kang, Z., & Maysami, R. C. (2008). Does Volatility Proxy Matter in Evaluating Volatility Forecasting Models? An Empirical Study.

Kang, J., & Park, H. J. (2008). The information content of net buying pressure: Evidence from the KOSPI 200 index option market. *Journal of Financial Markets*, 11(1), 36-56.

Latane, H. A., & Rendleman, R. J. (1976). Standard deviations of stock price ratios implied in option prices. *The Journal of Finance*, 31(2), 369-381.

Lamoureux, C. G., & Lastrapes, W. D. (1993). Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *The Review of Financial Studies*, 6(2), 293-326.

Lee, R. W. (2004). Implied volatility: Statics, dynamics, and probabilistic interpretation. *Recent advances in applied probability*, 241-268.

Lynch, P. E., & Zumbach, G. O. (2003). Market heterogeneities and the causal structure of volatility. *Quantitative Finance*, 3(4), 320.

McAleer, M., & Medeiros, M. C. (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics*, 147(1), 104-119.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2), 125-144.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Pictet, O. V., Olsen, R. B., & Ward, J. R. (1993). Fractals and intrinsic time: A challenge to econometricians. *Unpublished manuscript, Olsen & Associates, Zürich*, 130.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2), pp. 347–370.

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, (53), 61–65.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256.

Pena, I., Rubio, G., & Serna, G. (1999). Why do we smile? On the determinants of the implied volatility function. *Journal of Banking & Finance*, 23(8), 1151-1179.

Poon, S-H. 2005. A Practical Guide to Forecasting Financial Market Volatility. Chichester: John Wiley & Sons Ltd.

Poon, S. H., & Granger, C. W. (2003). Forecasting volatility in financial markets: A review. *Journal of economic literature*, 41(2), 478-539.

Press, S. J. (1967). A compound events model for security prices. *Journal of business*, 317-335.

Rogers, L. C. G., and Satchell, S. E. (1991). Estimating Variance from High, Low and Closing Prices. *Annals of Applied Probability*, 1, 504–512.

Rubinstein, M. (1994). Implied binomial trees. *The journal of finance*, 49(3), 771-818.

Shu, J., & Zhang, J. E. (2006). Testing range estimators of historical volatility. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 26(3), 297-313.

Taylor, S. J. (1986) Forecasting the Volatility of Currency Exchange Rates. *International Journal of Forecasting*, 3, 159–170.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical finance*, 4(2), 183-204.

Wong, A. H. S., & Heaney, R. A. (2017). Volatility smile and one-month foreign currency volatility forecasts. *Journal of Futures Markets*, 37(3), 286-312.

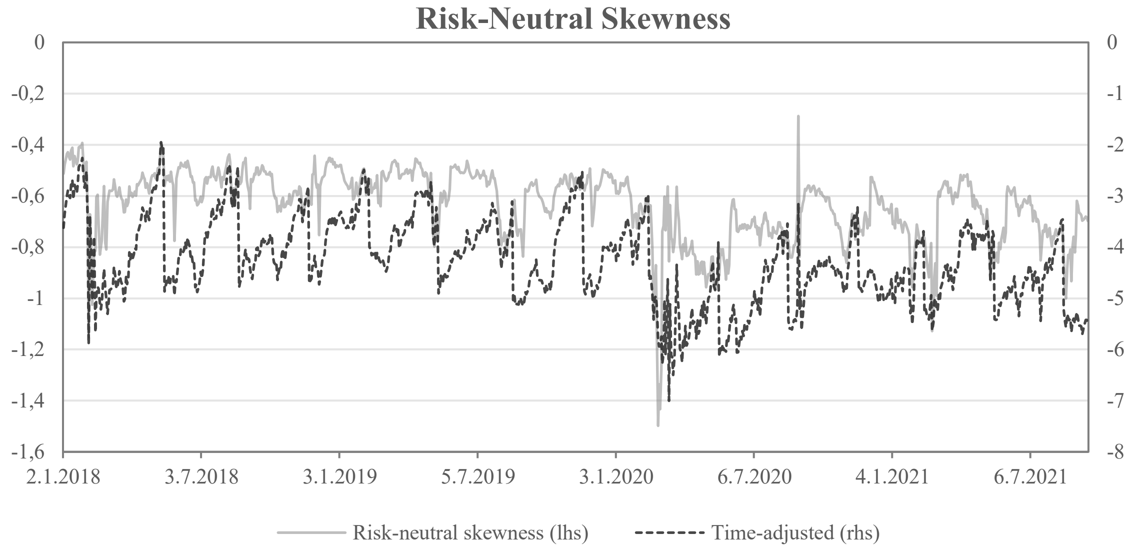
Yan, S. (2011). Jump risk, stock returns, and slope of implied volatility smile. *Journal of Financial Economics*, 99(1), 216-233.

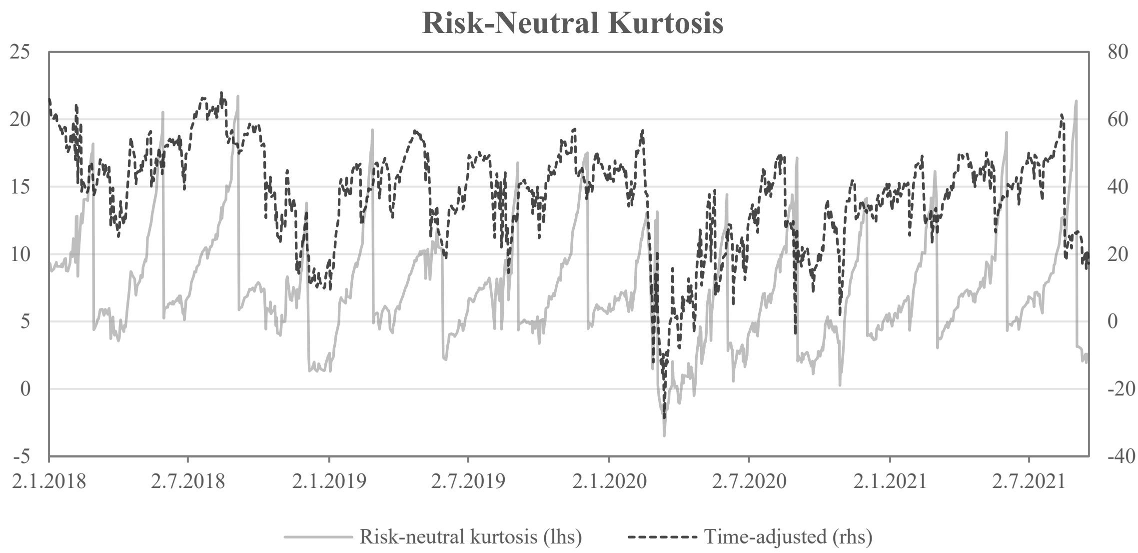
Yang, D., & Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *The Journal of Business*, 73(3), 477-492.

Xing, Y., Zhang, X., & Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns?. *Journal of Financial and Quantitative Analysis*, 45(3), 641-662.

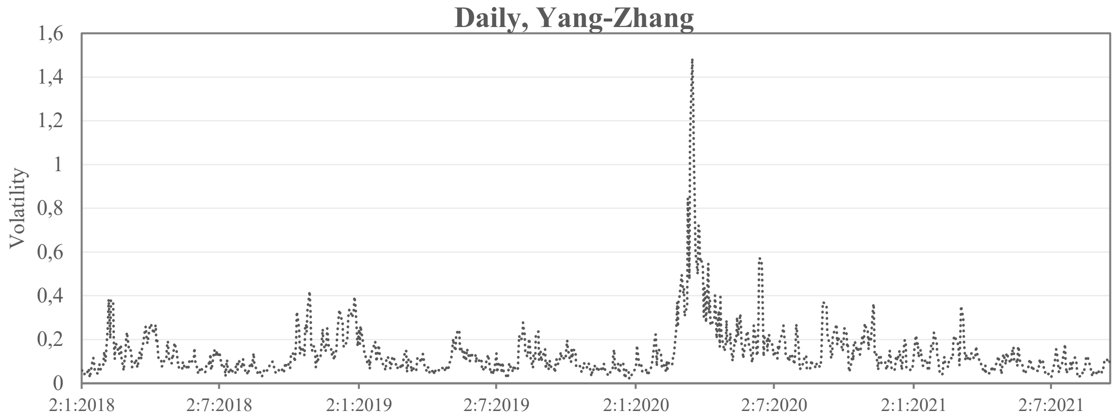
Xu, X., & Taylor, S. J. (1994). The term structure of volatility implied by foreign exchange options. *Journal of Financial and Quantitative Analysis*, 29(1), 57-74.

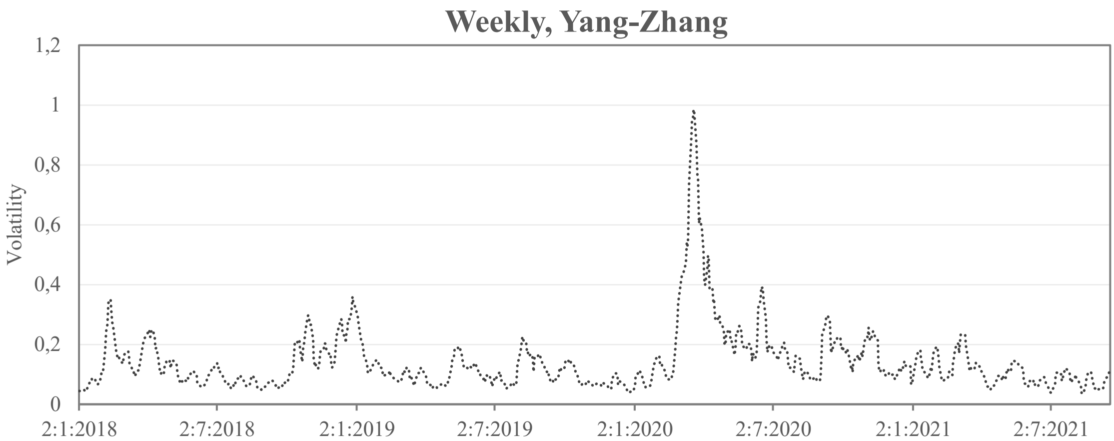
Appendix 1. Risk-neutral higher moments

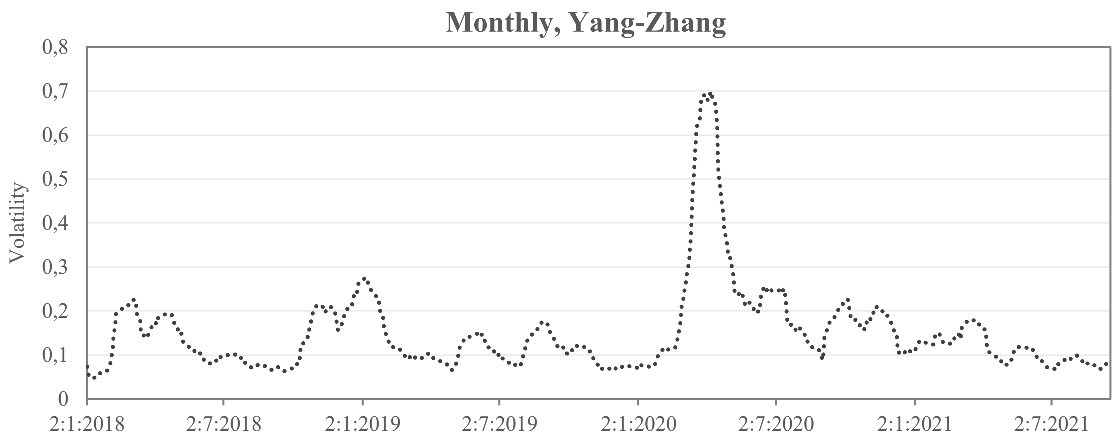


Appendix 2. Daily, weekly and monthly estimated Yang-Zhang volatilities



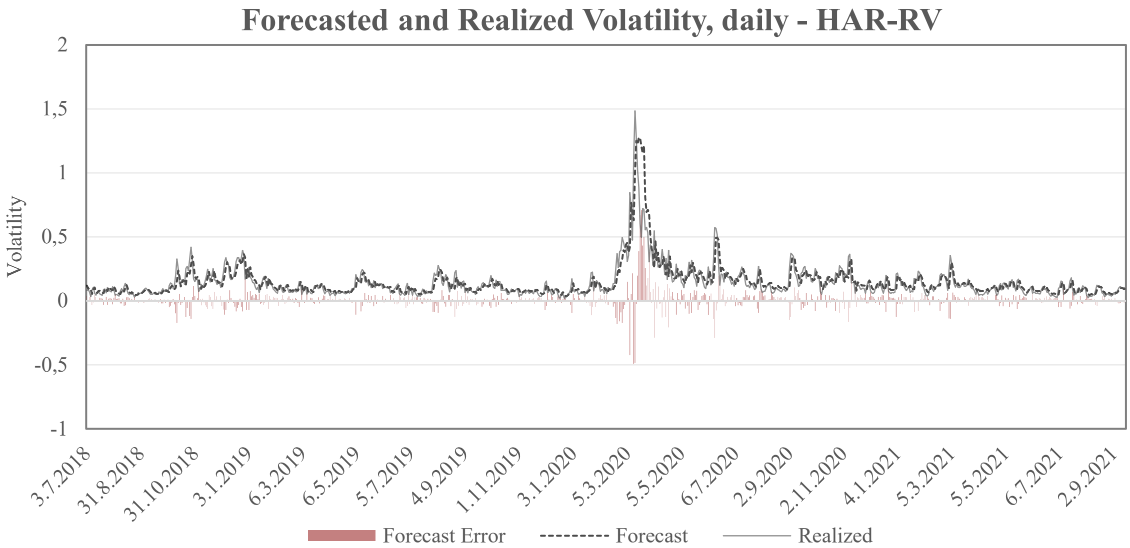



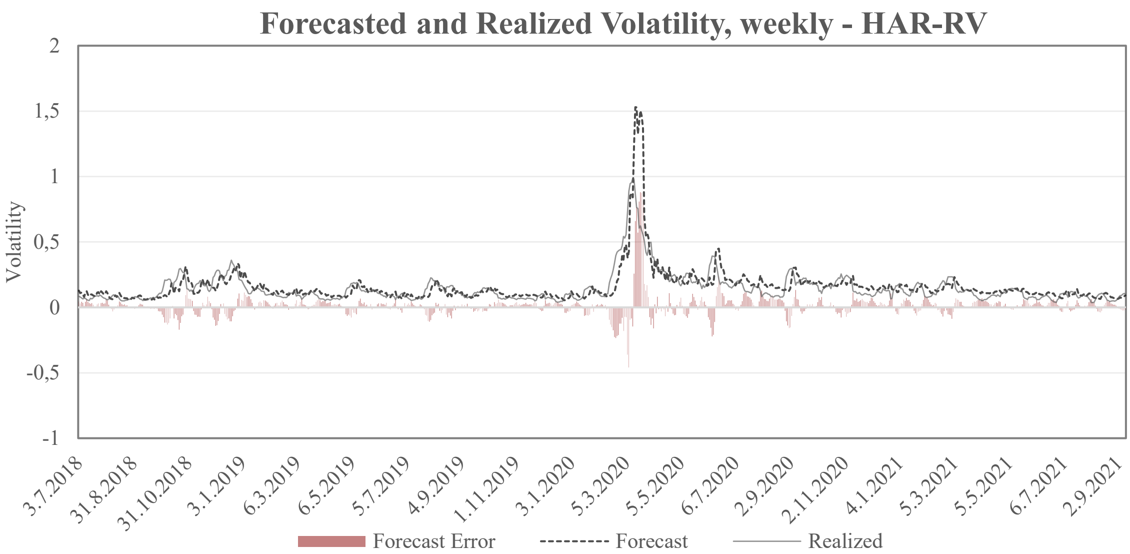
Appendix 3. Correlation matrix of used variables

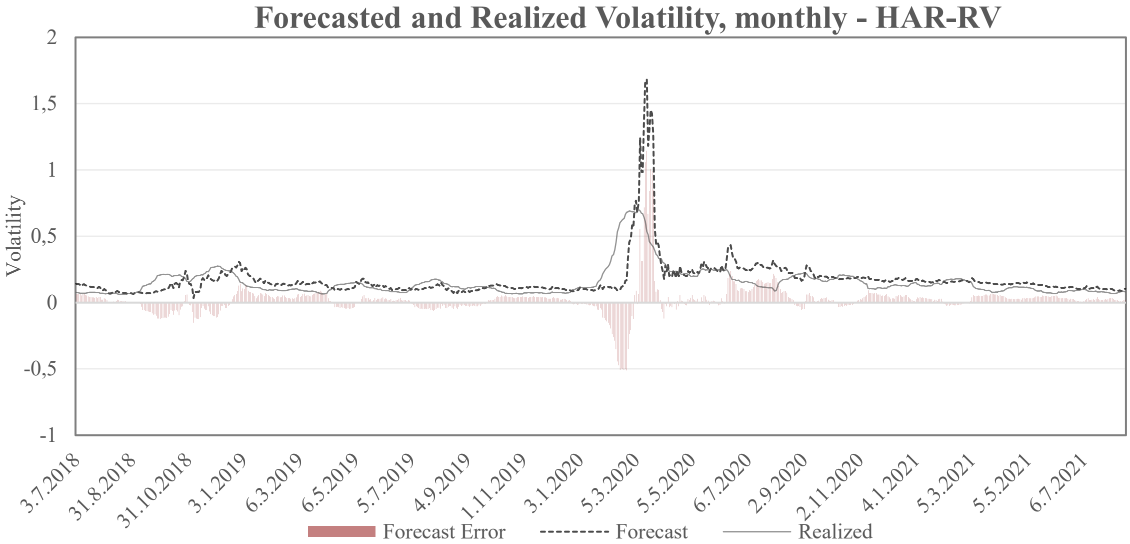
**Correlation matrix**

| | *σ, daily* | *σ, weekly* | *σ, monthly* | *SK* | *KU* | *σ, implied* |
|---|---|---|---|---|---|---|
| *σ, daily* | 1 | 0,89 | 0,70 | -0,39 | -0,65 | 0,77 |
| *σ, weekly* | 0,89 | 1 | 0,84 | -0,41 | -0,70 | 0,84 |
| *σ, monthly* | 0,70 | 0,84 | 1 | -0,40 | -0,69 | 0,83 |
| *SK* | -0,39 | -0,41 | -0,40 | 1 | 0,64 | -0,57 |
| *KU* | -0,65 | -0,70 | -0,69 | 0,64 | 1 | -0,81 |
| *σ, implied* | 0,77 | 0,84 | 0,83 | -0,57 | -0,81 | 1 |

Appendix 4. Volatility forecasts, HAR-RV

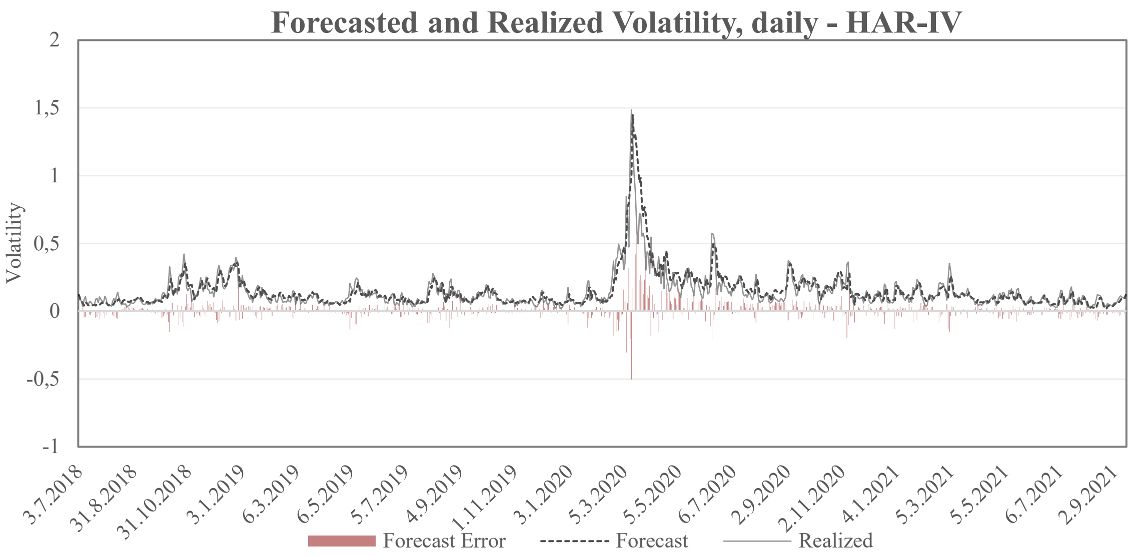


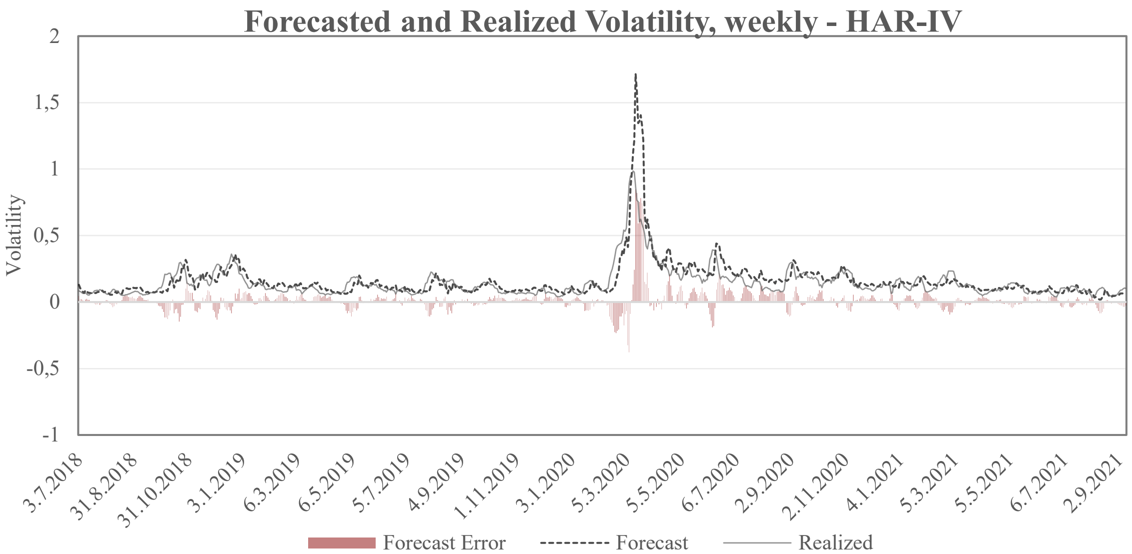



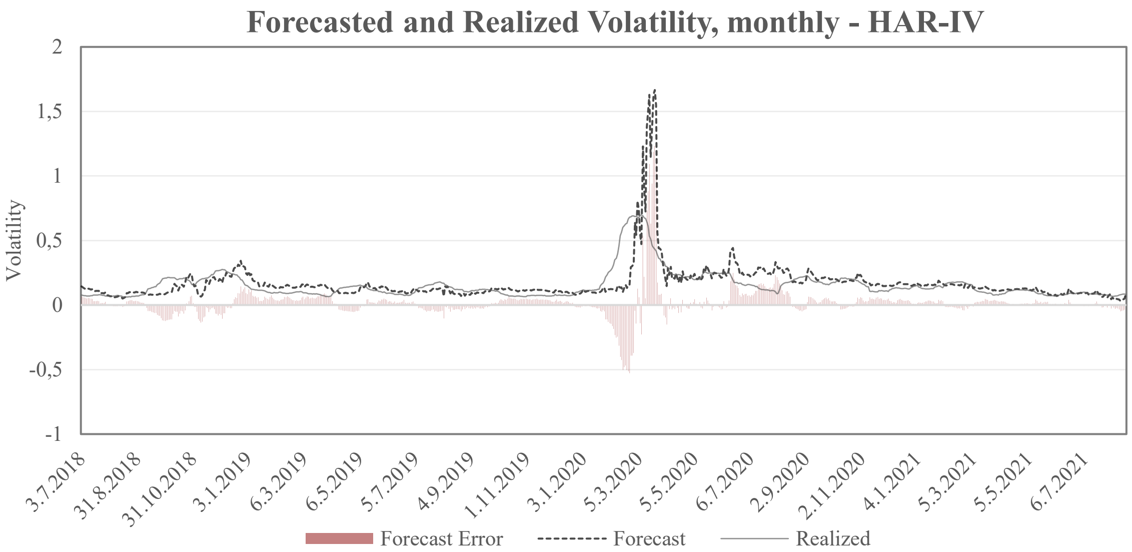


Forecasted and Realized Volatility, monthly - HAR-RV
Volatility
2
1,5
1
0,5
0
-0,5
-1
3.7.2018
31.8.2018
31.10.2018
3.1.2019
6.3.2019
6.5.2019
5.7.2019
4.9.2019
1.11.2019
3.1.2020
5.3.2020
5.5.2020
6.7.2020
2.9.2020
2.11.2020
4.1.2021
5.3.2021
5.5.2021
6.7.2021
Forecast Error
Forecast
Realized

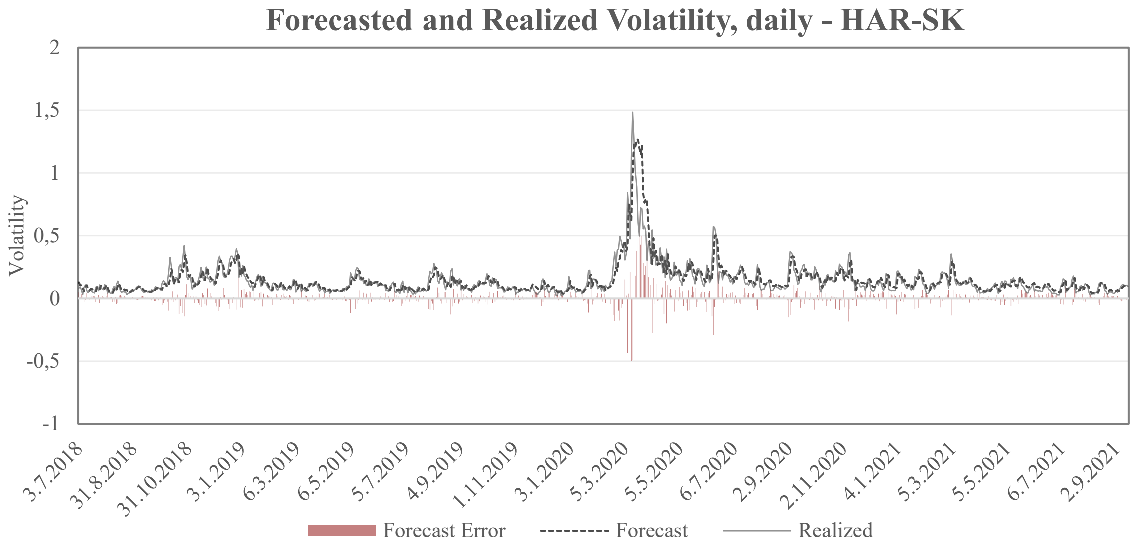
Appendix 5. Volatility forecasts, HAR-IV

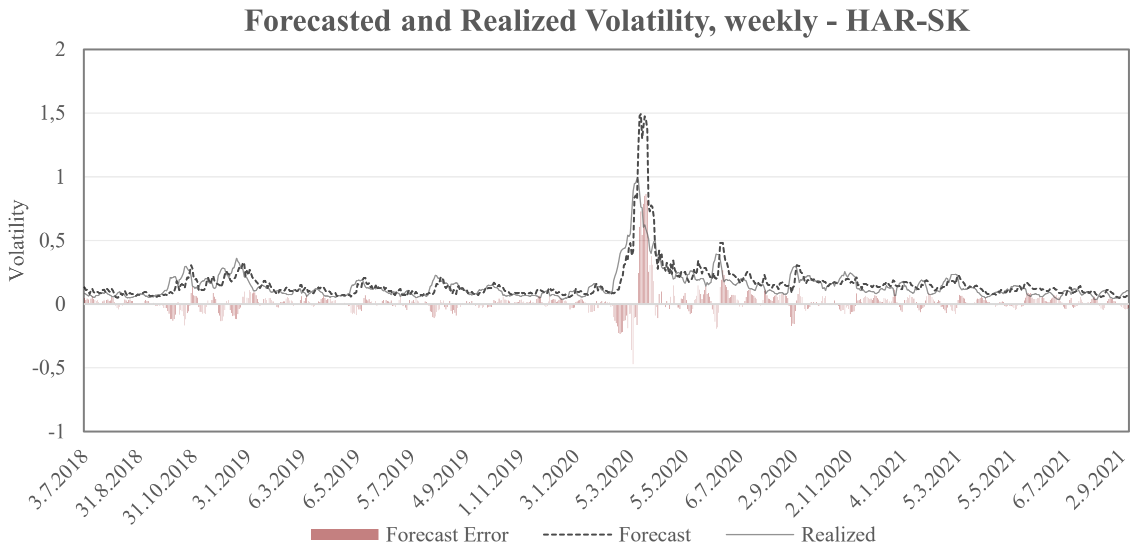


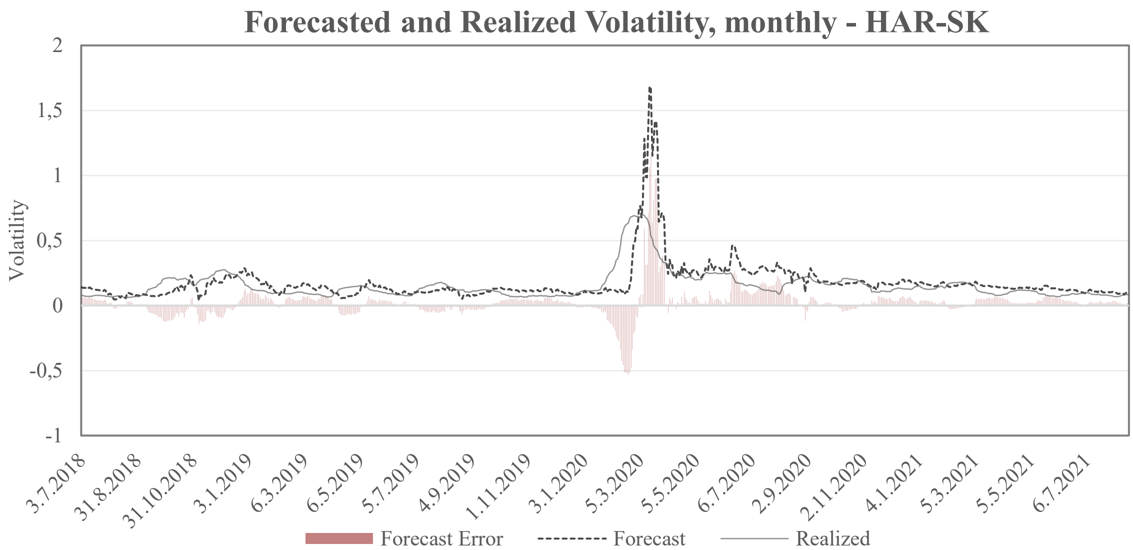



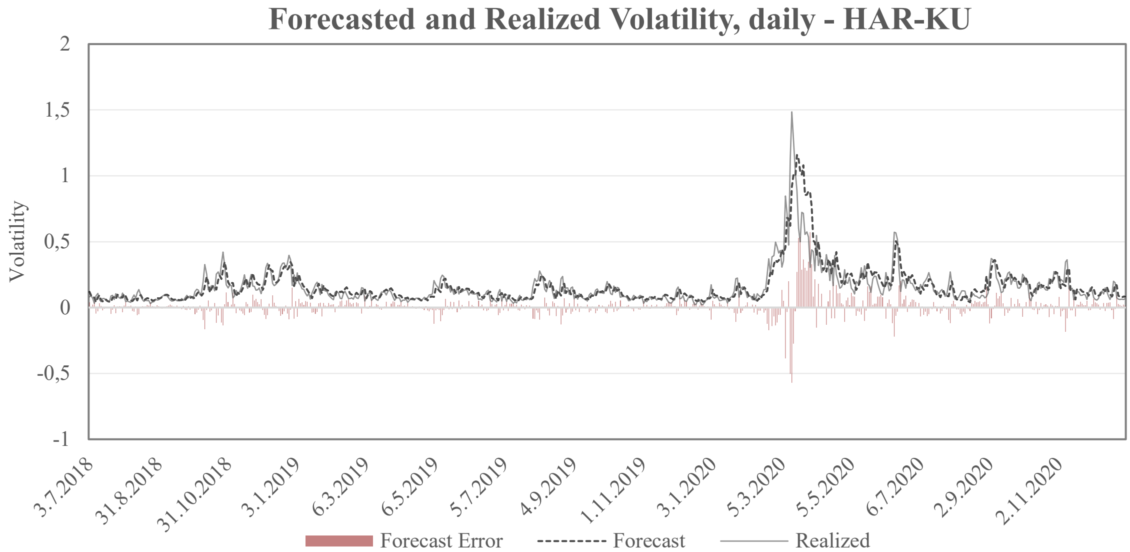
Appendix 6. Volatility forecasts, HAR-SK

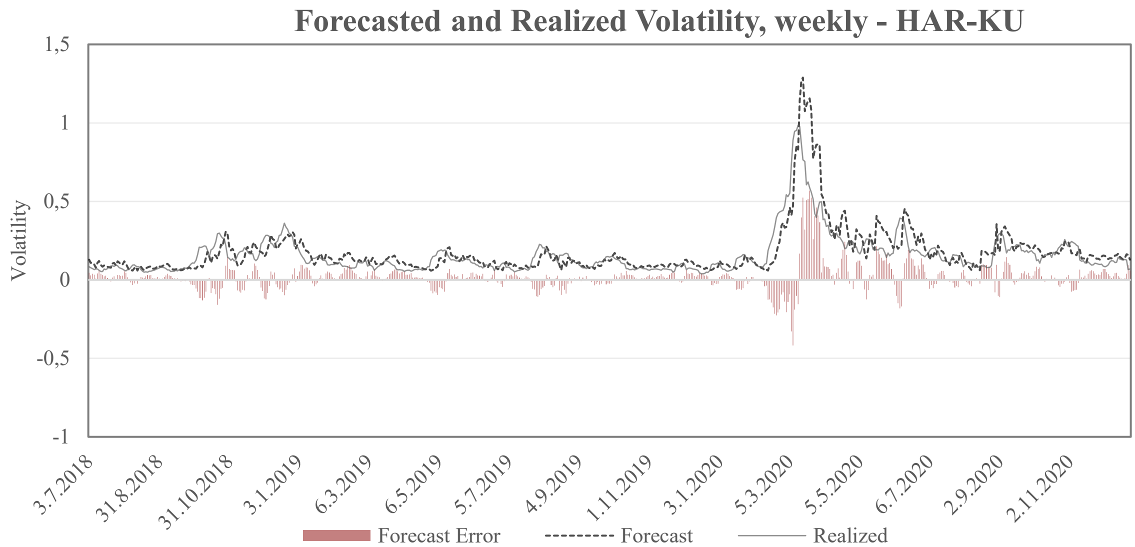


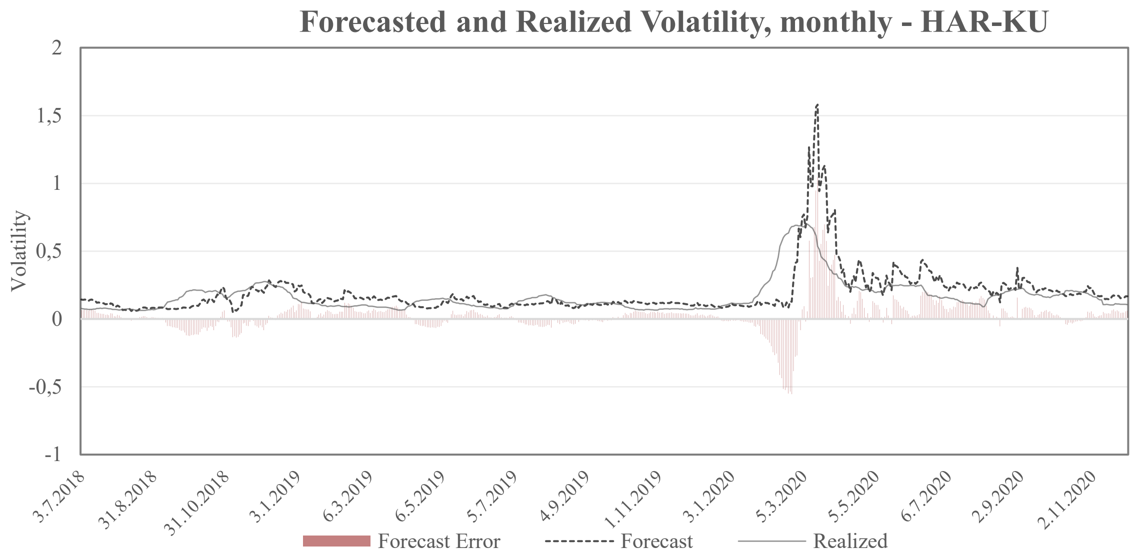



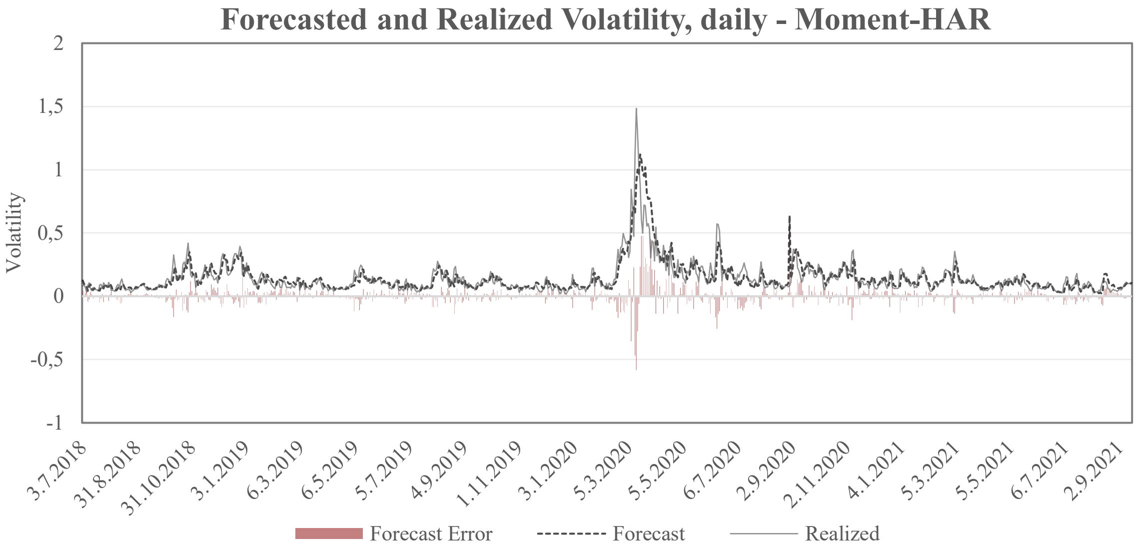
Appendix 7. Volatility forecasts, HAR-KU

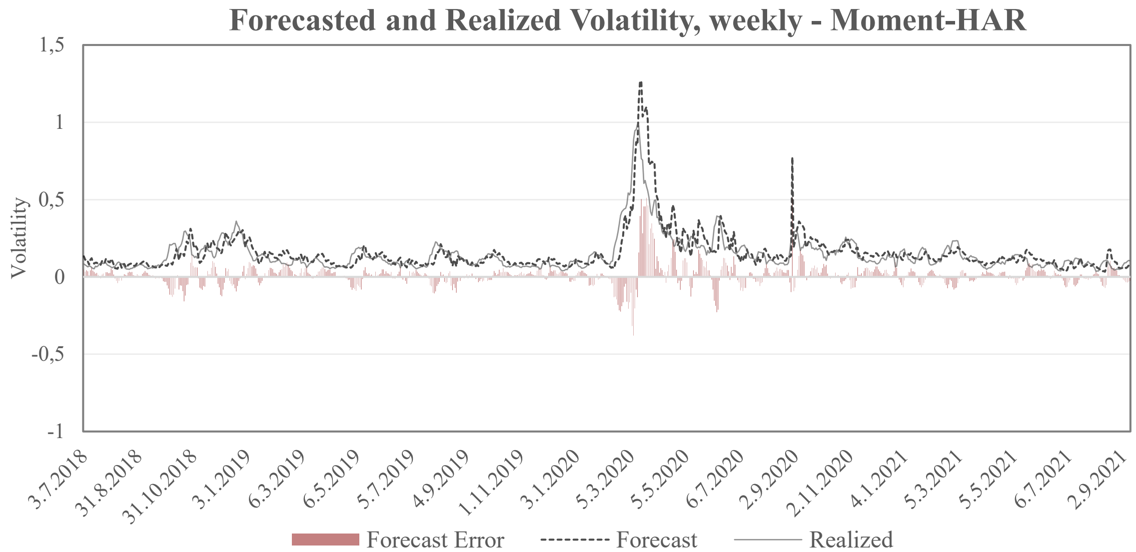


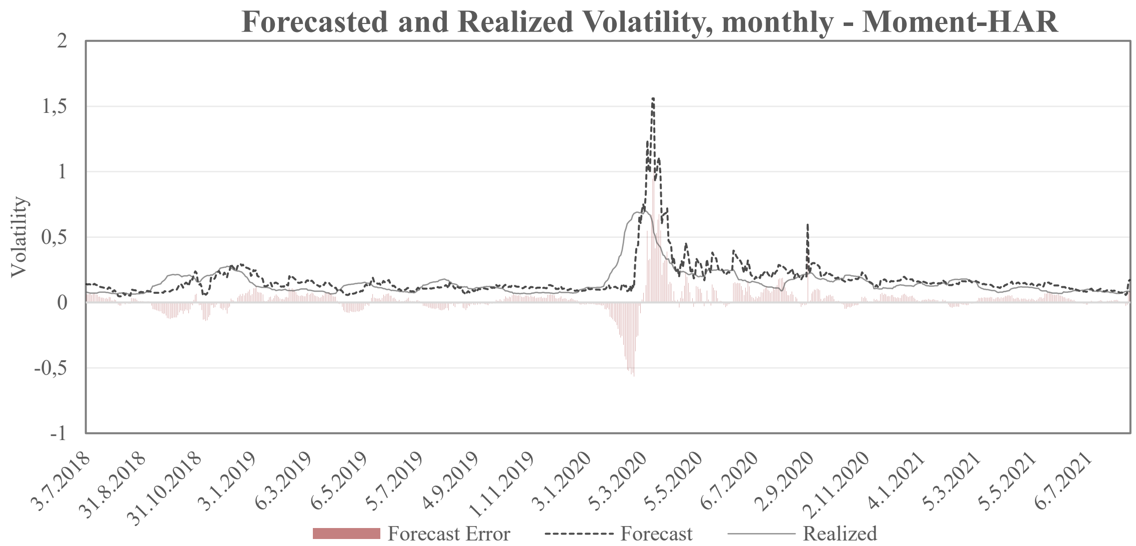



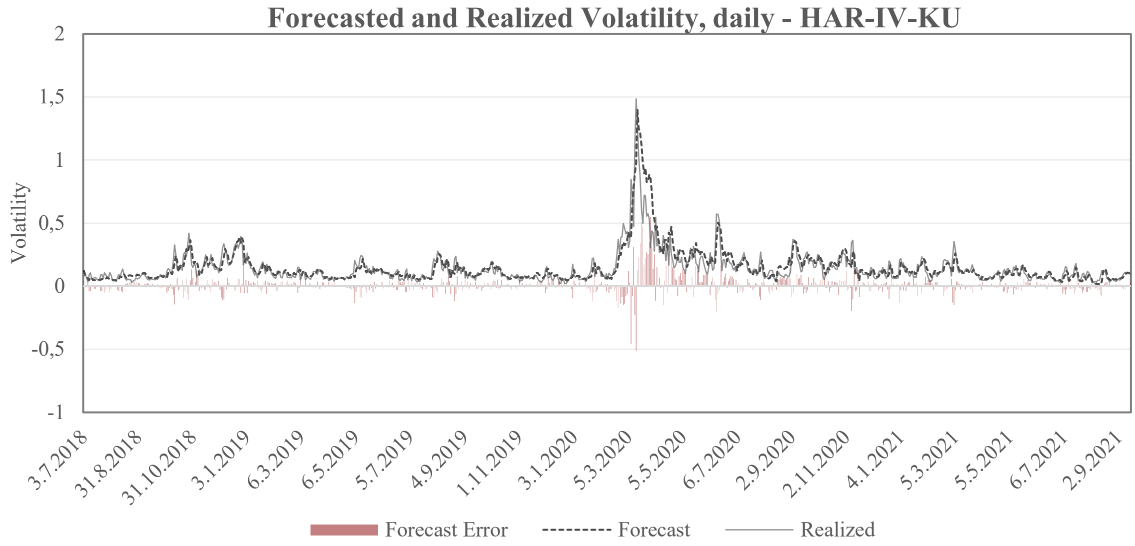
Appendix 8. Volatility forecasts, Moment-HAR

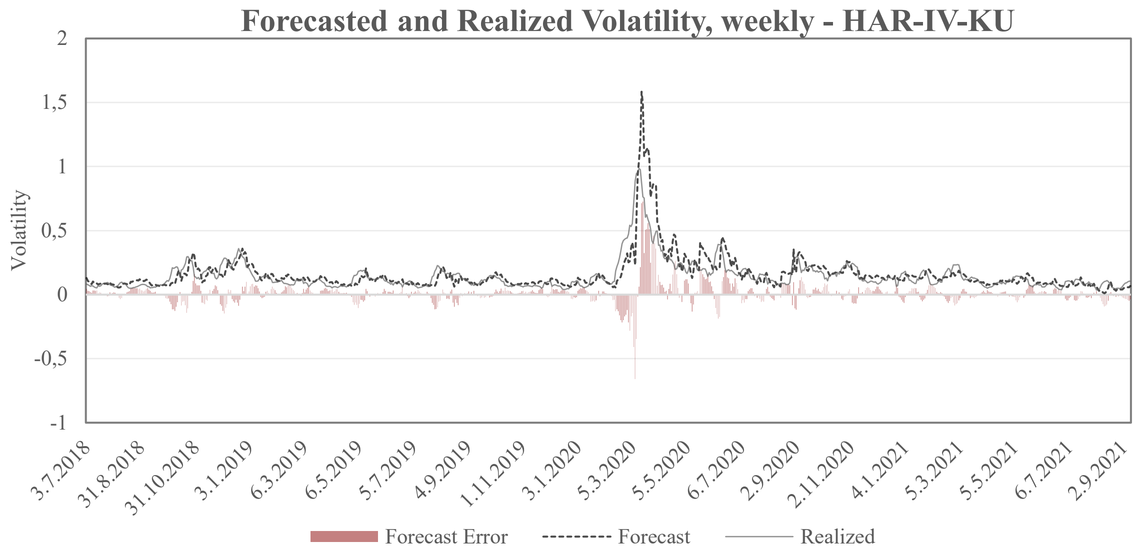


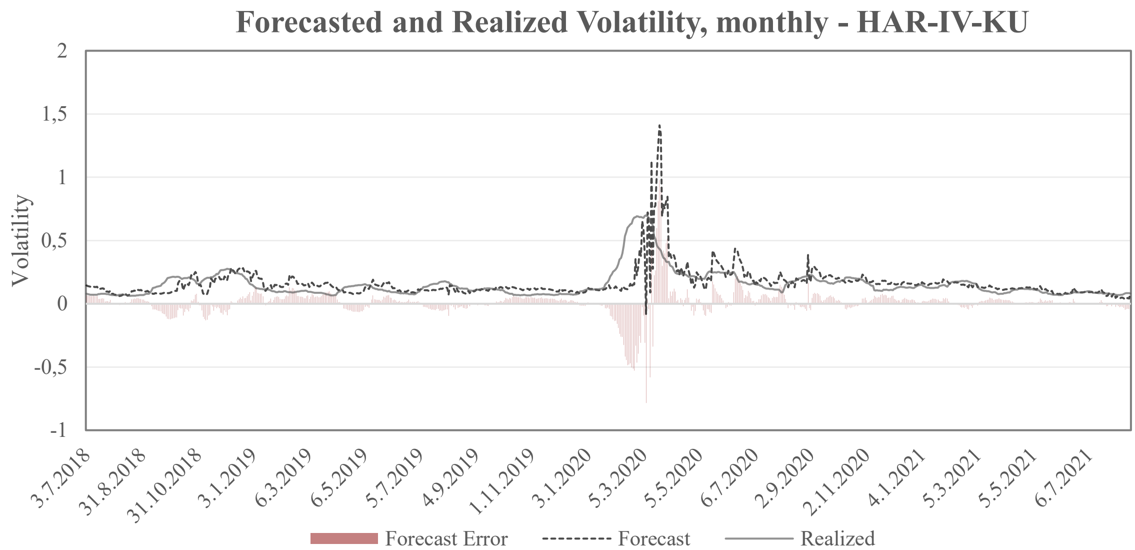



Appendix 9. Volatility forecasts, HAR-IV-KU



Appendix 10. Coefficient overview, HAR-RV

**Coeffiecent overview over 163 recalibrations, HAR-RV, daily forecasts**

| | **Daily volatility** | **Weekly volatility** | **Monthly volatility** |
|---|---|---|---|
| *Maximum* | 0,832 | 0,547 | 1,273 |
| *(p-value)* | *(0)* | *(0,003)* | *(0,003)* |
| *Minimum* | 0,260 | -0,174 | -0,296 |
| *(p-value)* | *(0,031)* | *(0,187)* | *(0,07)* |
| *Median* | 0,683 | 0,139 | 0,000 |
| *(p-value)* | *(0)* | *(0,64)* | *(0,964)* |
| *Mean* | 0,671 | 0,154 | 0,003 |
| *No. of significant coefficients* | 162 | 51 | 1 |
| *%* | 99 % | 31 % | 1 % |

**Coeffiecent overview over 162 recalibrations, HAR-RV, weekly forecasts**

| | **Daily volatility** | **Weekly volatility** | **Monthly volatility** |
|---|---|---|---|
| *Maximum* | 0,822 | 0,763 | 1,615 |
| *(p-value)* | *(0)* | *(0,007)* | *(0)* |
| *Minimum* | -0,280 | -0,170 | -0,581 |
| *(p-value)* | *(0,019)* | *(0,419)* | *(0)* |
| *Median* | 0,362 | 0,229 | -0,074 |
| *(p-value)* | *(0,027)* | *(0,204)* | *(0,101)* |
| *Mean* | 0,376 | 0,263 | -0,051 |
| *No. of significant coefficients* | 161 | 79 | 63 |
| *%* | 99 % | 48 % | 39 % |

**Coeffiecent overview over 157 recalibrations, HAR-RV, monthly forecasts**

| | **Daily volatility** | **Weekly volatility** | **Monthly volatility** |
|---|---|---|---|
| *Maximum* | 0,858 | 0,937 | 1,729 |
| *(p-value)* | *(0,016)* | *(0)* | *(0,008)* |
| *Minimum* | -0,499 | -0,103 | -0,946 |
| *(p-value)* | *(0,024)* | *(0,374)* | *(0)* |
| *Median* | 0,170 | 0,132 | -0,135 |
| *(p-value)* | *(0,092)* | *(0,882)* | *(0,034)* |
| *Mean* | 0,181 | 0,180 | -0,113 |
| *No. of significant coefficients* | 74 | 38 | 79 |
| *%* | 45 % | 23 % | 48 % |

Appendix 11. Coefficient overview, HAR-IV

**Coeffiecent overview over 163 recalibrations, HAR-IV, daily forecasts**

| | **Daily volatility** | **Weekly volatility** | **Monthly volatility** | **Implied volatility** |
|---|---|---|---|---|
| *Maximum* | 0,751 | 0,395 | 0,576 | 2,344 |
| *(p-value)* | *(0)* | *(0,007)* | *(0,172)* | *(0)* |
| *Minimum* | 0,246 | -0,563 | -0,761 | 0,101 |
| *(p-value)* | *(0,006)* | *(0,046)* | *(0)* | *(0,757)* |
| *Median* | 0,580 | -0,039 | -0,195 | 0,855 |
| *(p-value)* | *(0)* | *(0,57)* | *(0,146)* | *(0,008)* |
| *Mean* | 0,558 | 0,005 | -0,209 | 0,980 |
| *No. of significant coefficients* | 162 | 18 | 67 | 157 |
| *%* | 99 % | 11 % | 41 % | 96 % |

**Coeffiecent overview over 162 recalibrations, HAR-IV, weekly forecasts**

| | **Daily volatility** | **Weekly volatility** | **Monthly volatility** | **Implied volatility** |
|---|---|---|---|---|
| *Maximum* | 0,748 | 0,556 | 1,161 | 2,107 |
| *(p-value)* | *(0)* | *(0,049)* | *(0,009)* | *(0)* |
| *Minimum* | -0,289 | -0,388 | -0,912 | -0,019 |
| *(p-value)* | *(0,013)* | *(0,006)* | *(0)* | *(0,943)* |
| *Median* | 0,257 | 0,119 | -0,285 | 0,942 |
| *(p-value)* | *(0,017)* | *(0,637)* | *(0,017)* | *(0)* |
| *Mean* | 0,255 | 0,111 | -0,254 | 1,003 |
| *No. of significant coefficients* | 135 | 56 | 105 | 135 |
| *%* | 83 % | 34 % | 64 % | 83 % |

**Coeffiecent overview over 157 recalibrations, HAR-IV, monthly forecasts**

| | **Daily volatility** | **Weekly volatility** | **Monthly volatility** | **Implied volatility** |
|---|---|---|---|---|
| *Maximum* | 1,233 | 1,215 | 2,337 | 1,851 |
| *(p-value)* | *(0)* | *(0,02)* | *(0,004)* | *(0)* |
| *Minimum* | -0,483 | -0,315 | -0,989 | -3,529 |
| *(p-value)* | *(0,024)* | *(0,003)* | *(0)* | *(0)* |
| *Median* | 0,118 | 0,073 | -0,236 | 0,628 |
| *(p-value)* | *(0,42)* | *(0,753)* | *(0,03)* | *(0)* |
| *Mean* | 0,119 | 0,140 | -0,158 | 0,358 |
| *No. of significant coefficients* | 49 | 44 | 82 | 109 |
| *%* | 30 % | 27 % | 50 % | 67 % |

Appendix 12. Coefficient overview, HAR-SK

**Coeffiecent overview over 163 recalibrations, HAR-SK, daily forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral skewness |
|---|---|---|---|---|
| *Maximum* | 0,818 | 0,517 | 1,422 | 0,008 |
| *(p-value)* | *(0)* | *(0,007)* | *(0,002)* | *(0,159)* |
| *Minimum* | 0,265 | -0,260 | -0,340 | -0,023 |
| *(p-value)* | *(0,028)* | *(0,411)* | *(0,061)* | *(0,001)* |
| *Median* | 0,671 | 0,152 | 0,011 | -0,008 |
| *(p-value)* | *(0,028)* | *(0,689)* | *(0,737)* | *(0,257)* |
| *Mean* | 0,654 | 0,150 | 0,004 | -0,008 |
| *No. of significant coefficients* | 162 | 52 | 3 | 18 |
| *%* | 99 % | 32 % | 2 % | 11 % |

**Coeffiecent overview over 162 recalibrations, HAR-SK, weekly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral skewness |
|---|---|---|---|---|
| *Maximum* | 0,795 | 0,576 | 2,089 | 0,016 |
| *(p-value)* | *(0)* | *(0)* | *(0)* | *(0)* |
| *Minimum* | -0,267 | -0,219 | -0,643 | -0,054 |
| *(p-value)* | *(0,018)* | *(0,3)* | *(0)* | *(0)* |
| *Median* | 0,340 | 0,250 | -0,066 | -0,011 |
| *(p-value)* | *(0,009)* | *(0,235)* | *(0,183)* | *(0,059)* |
| *Mean* | 0,353 | 0,252 | -0,044 | -0,014 |
| *No. of significant coefficients* | 161 | 79 | 67 | 64 |
| *%* | 99 % | 48 % | 41 % | 39 % |

**Coeffiecent overview over 157 recalibrations, HAR-SK, monthly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral skewness |
|---|---|---|---|---|
| *Maximum* | 0,926 | 0,950 | 2,046 | 0,042 |
| *(p-value)* | *(0,01)* | *(0)* | *(0,011)* | *(0,002)* |
| *Minimum* | -0,484 | -0,260 | -0,810 | -0,068 |
| *(p-value)* | *(0,025)* | *(0,624)* | *(0)* | *(0,002)* |
| *Median* | 0,159 | 0,138 | -0,117 | -0,011 |
| *(p-value)* | *(0,318)* | *(0,811)* | *(0,048)* | *(0,058)* |
| *Mean* | 0,159 | 0,173 | -0,076 | -0,013 |
| *No. of significant coefficients* | 49 | 39 | 88 | 84 |
| *%* | 30 % | 24 % | 54 % | 52 % |

Appendix 13. Coefficient overview, HAR-KU

**Coeffiecent overview over 163 recalibrations, HAR-KU, daily forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral kurtosis |
|---|---|---|---|---|
| *Maximum* | 0,784 | 0,438 | 1,294 | 0,000 |
| *(p-value)* | *(0)* | *(0,004)* | *(0,002)* | *(0,358)* |
| *Minimum* | 0,238 | -0,631 | -0,686 | -0,005 |
| *(p-value)* | *(0,021)* | *(0,02)* | *(0,015)* | *(0)* |
| *Median* | 0,613 | 0,048 | -0,089 | -0,002 |
| *(p-value)* | *(0)* | *(0,777)* | *(0,083)* | *(0)* |
| *Mean* | 0,608 | 0,091 | -0,103 | -0,002 |
| *No. of significant coefficients* | 162 | 34 | 34 | 130 |
| *%* | 99 % | 21 % | 21 % | 80 % |

**Coeffiecent overview over 162 recalibrations, HAR-KU, weekly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral kurtosis |
|---|---|---|---|---|
| *Maximum* | 0,584 | 0,532 | 1,658 | 0,001 |
| *(p-value)* | *(0)* | *(0)* | *(0)* | *(0,007)* |
| *Minimum* | -0,177 | -0,505 | -0,981 | -0,009 |
| *(p-value)* | *(0,049)* | *(0,023)* | *(0)* | *(0)* |
| *Median* | 0,252 | 0,211 | -0,172 | -0,002 |
| *(p-value)* | *(0,018)* | *(0,799)* | *(0,005)* | *#N/A* |
| *Mean* | 0,292 | 0,179 | -0,178 | -0,003 |
| *No. of significant coefficients* | 137 | 73 | 73 | 129 |
| *%* | 84 % | 45 % | 45 % | 79 % |

**Coeffiecent overview over 157 recalibrations, HAR-KU, monthly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral kurtosis |
|---|---|---|---|---|
| *Maximum* | 1,138 | 1,011 | 1,710 | 0,007 |
| *(p-value)* | *(0,003)* | *(0)* | *(0,009)* | *(0)* |
| *Minimum* | -0,399 | -0,630 | -1,043 | -0,009 |
| *(p-value)* | *(0,056)* | *(0,166)* | *(0)* | *(0)* |
| *Median* | 0,124 | 0,091 | -0,172 | -0,002 |
| *(p-value)* | *(0,45)* | *(0,885)* | *(0,032)* | *(0,002)* |
| *Mean* | 0,128 | 0,103 | -0,160 | -0,002 |
| *No. of significant coefficients* | 38 | 40 | 79 | 120 |
| *%* | 23 % | 25 % | 48 % | 74 % |

Appendix 14. Coefficient overview, Moment-HAR

**Coeffiecent overview over 163 recalibrations,
Moment-HAR, daily forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral skewness | Risk-neutral kurtosis |
|---|---|---|---|---|---|
| *Maximum* | 0,784 | 0,526 | 1,025 | 0,116 | 0,000 |
| *(p-value)* | *(0)* | *(0,004)* | *(0,019)* | *(0)* | *(0,575)* |
| *Minimum* | 0,229 | -0,393 | -0,689 | -0,010 | -0,009 |
| *(p-value)* | *(0,021)* | *(0,153)* | *(0,013)* | *(0,271)* | *(0)* |
| *Median* | 0,608 | 0,039 | -0,131 | 0,008 | -0,002 |
| *(p-value)* | *(0)* | *(0,943)* | *(0,04)* | *(0,442)* | *(0,024)* |
| *Mean* | 0,599 | 0,089 | -0,140 | 0,017 | -0,003 |
| *No. of significant coefficients* | 162 | 39 | 38 | 37 | 120 |
| *%* | 99 % | 24 % | 23 % | 23 % | 74 % |

**Coeffiecent overview over 162 recalibrations,
Moment-HAR, weekly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral skewness | Risk-neutral kurtosis |
|---|---|---|---|---|---|
| *Maximum* | 0,564 | 0,539 | 1,382 | 0,122 | 0,001 |
| *(p-value)* | *(0)* | *(0)* | *(0)* | *(0)* | *(0,017)* |
| *Minimum* | -0,162 | -0,317 | -0,982 | -0,018 | -0,011 |
| *(p-value)* | *(0,066)* | *(0,116)* | *(0)* | *(0,106)* | *(0)* |
| *Median* | 0,240 | 0,170 | -0,225 | 0,009 | -0,003 |
| *(p-value)* | *(0,046)* | *(0,216)* | *(0,001)* | *(0,24)* | *(0)* |
| *Mean* | 0,282 | 0,175 | -0,217 | 0,017 | -0,003 |
| *No. of significant coefficients* | 131 | 68 | 79 | 56 | 124 |
| *%* | 80 % | 42 % | 48 % | 34 % | 76 % |

**Coeffiecent overview over 157 recalibrations,
Moment-HAR, monthly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Risk-neutral skewness | Risk-neutral kurtosis |
|---|---|---|---|---|---|
| *Maximum* | 1,104 | 1,005 | 1,619 | 0,137 | 0,008 |
| *(p-value)* | *(0,007)* | *(0)* | *(0,014)* | *(0)* | *(0,001)* |
| *Minimum* | -0,393 | -0,457 | -1,139 | -0,050 | -0,011 |
| *(p-value)* | *(0,061)* | *(0,338)* | *(0)* | *(0)* | *(0)* |
| *Median* | 0,102 | 0,084 | -0,184 | 0,005 | -0,002 |
| *(p-value)* | *(0,434)* | *(0,909)* | *(0,059)* | *(0,776)* | *(0)* |
| *Mean* | 0,117 | 0,110 | -0,177 | 0,011 | -0,002 |
| *No. of significant coefficients* | 33 | 42 | 80 | 63 | 104 |
| *%* | 20 % | 26 % | 49 % | 39 % | 64 % |

Appendix 15. Coefficient overview, HAR-IV-KU

**Coeffiecent overview over 163 recalibrations, HAR-IV-KU, daily forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Implied volatility | Risk-neutral kurtosis |
|---|---|---|---|---|---|
| *Maximum* | 0,747 | 0,363 | 0,814 | 2,660 | 0,003 |
| *(p-value)* | *(0)* | *(0,014)* | *(0,019)* | *(0)* | *(0,005)* |
| *Minimum* | 0,214 | -0,853 | -0,564 | -0,783 | -0,004 |
| *(p-value)* | *(0,017)* | *(0,002)* | *(0,01)* | *(0,122)* | *(0)* |
| *Median* | 0,551 | -0,042 | -0,209 | 1,006 | 0,001 |
| *(p-value)* | *(0)* | *(0,741)* | *(0,044)* | *(0,003)* | *(0,842)* |
| *Mean* | 0,529 | 0,006 | -0,206 | 1,111 | 0,000 |
| *No. of significant coefficients* | 162 | 16 | 68 | 115 | 68 |
| *%* | 99 % | 10 % | 42 % | 71 % | 42 % |

**Coeffiecent overview over 162 recalibrations, HAR-IV-KU, weekly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Implied volatility | Risk-neutral kurtosis |
|---|---|---|---|---|---|
| *Maximum* | 0,525 | 0,591 | 1,607 | 3,358 | 0,003 |
| *(p-value)* | *(0)* | *(0)* | *(0)* | *(0)* | *(0)* |
| *Minimum* | -0,179 | -0,586 | -0,914 | -2,848 | -0,009 |
| *(p-value)* | *(0,048)* | *(0,011)* | *(0,001)* | *(0)* | *(0)* |
| *Median* | 0,220 | 0,130 | -0,317 | 0,701 | -0,001 |
| *(p-value)* | *(0,201)* | *(0,771)* | *(0,001)* | *(0,003)* | *(0,014)* |
| *Mean* | 0,218 | 0,128 | -0,258 | 0,805 | -0,001 |
| *No. of significant coefficients* | 114 | 65 | 108 | 89 | 86 |
| *%* | 70 % | 40 % | 66 % | 55 % | 53 % |

**Coeffiecent overview over 157 recalibrations, HAR-IV-KU, monthly forecasts**

| | Daily volatility | Weekly volatility | Monthly volatility | Implied volatility | Risk-neutral kurtosis |
|---|---|---|---|---|---|
| *Maximum* | 0,788 | 1,112 | 2,956 | 2,971 | 0,003 |
| *(p-value)* | *(0,028)* | *(0)* | *(0)* | *(0)* | *(0)* |
| *Minimum* | -0,331 | -0,357 | -0,961 | -8,582 | -0,015 |
| *(p-value)* | *(0,081)* | *(0,016)* | *(0)* | *(0)* | *(0)* |
| *Median* | 0,116 | 0,039 | -0,243 | 0,096 | 0,000 |
| *(p-value)* | *(0,475)* | *(0,801)* | *(0,006)* | *(0,514)* | *(0,143)* |
| *Mean* | 0,101 | 0,138 | -0,127 | -0,059 | -0,002 |
| *No. of significant coefficients* | 47 | 48 | 93 | 106 | 91 |
| *%* | 29 % | 29 % | 57 % | 65 % | 56 % |