



LUND UNIVERSITY
School of Economics and Management

Forecasting Volatility:
Evidence From The Swiss Stock Market

DEGREE PROJECT I

Spring 2010

Authors

Dennis A. Huber

Jacob N. Persson

Supervisor

Hossein Asgharian

Table of content

Table of content	i
Abstract	iii
1 Introduction	1
1.1 Background	1
1.2 Earlier research	2
1.3 Problem description	3
1.4 Purpose.....	4
1.5 Problem delimitation.....	4
1.6 Disposition	4
2 Theoretical background.....	5
2.1 Types of volatility	5
2.1.1 <i>Historical volatility</i>	5
2.1.2 <i>Realized volatility</i>	5
2.1.3 <i>Implied volatility</i>	6
2.2 Characteristics of volatility	7
2.2.1 <i>Volatility clustering</i>	7
2.2.2 <i>Leverage effect</i>	8
2.2.3 <i>Volatility smile</i>	8
2.2.4 <i>Distribution characteristics</i>	9
2.3 Modelling volatility	10
2.3.1 <i>Naïve models</i>	10
2.3.2 <i>Autoregressive Moving Average (ARMA)</i>	10
2.3.3 <i>Autoregressive Integrated Moving Average (ARIMA)</i>	12
2.3.4 <i>Autoregressive Conditional Heteroscedasticity (ARCH)</i>	13
2.3.5 <i>Generalized ARCH (GARCH)</i>	13
2.3.6 <i>Exponential GARCH (EGARCH)</i>	14
2.4 Quality evaluation of forecasts	15
2.4.1 <i>Root Mean Square Error (RMSE)</i>	15
2.4.2 <i>Mean Absolute Error (MAE)</i>	15
2.4.3 <i>Mean Absolute Percentage Error (MAPE)</i>	16
2.4.4 <i>Adjusted Mean Absolute Percentage Error (AMAPE)</i>	16
2.4.5 <i>Theil-U statistic</i>	16
2.4.6 <i>LINEX</i>	17
3 Methodology.....	17
3.1 Matching time horizons	18

3.2	Forecasting realized volatility.....	18
3.3	Investigating the forecast quality.....	20
3.4	Predictive power and information content.....	20
3.5	Additional information in VSMI.....	21
3.6	Choice of method.....	23
3.7	Criticism of the method.....	23
4	Data description.....	24
4.1	General data description.....	24
4.2	Swiss Market Index (SMI).....	24
4.3	VSMI as a volatility index.....	25
4.4	Descriptive Statistics.....	27
5	Empirical results.....	29
5.1	Descriptive statistics of rolling estimation period.....	29
5.2	Model estimation.....	30
5.3	Descriptive statistics of generated forecasts.....	33
5.4	Forecast evaluation.....	34
5.5	Predictive Power.....	36
5.6	Information Content.....	37
5.7	Test for additional information in VSMI.....	38
6	Conclusion.....	40
7	Appendix.....	iv
	References.....	xiii
	List of figures.....	xvii
	List of tables.....	xviii
	List of used abbreviations.....	xix

Abstract

This thesis focuses on the question whether the volatility index of the Swiss Market Index (VSMI) is an adequate predictor of the future realized volatility. Furthermore, the thesis searches for evidence of incremental information within the VSMI in relation to a set of model based forecasts (MBFs) which could describe future realized volatility. First a theoretical framework, corresponding to the used methodology, is presented. Then the VSMI is assessed in terms of error accuracy, predictive power as well as information content in comparison to generally used models. Additionally, a generalized method of moments (GMM) optimization is used to investigate whether there is additional information contained in the volatility index. The results indicate that the VSMI is the best predictor for future volatility and there is evidence that this volatility index holds incremental information to the extent of information implied in the used set of MBFs.

1 Introduction

1.1 Background

Since the ground-breaking work of Markowitz (1952) on portfolio selection, expected return and volatility are considered the two fundamental variables in modern financial theory. Measured by the standard deviation or variance of returns, volatility is a simple measure of the total risk of financial assets.¹ Therefore, volatility is a key element in the decision making process of investors and traders as well as input to a wide range of financial models.²

One such model is option pricing introduced by Black and Scholes (1973). It provides an easy way to calculate the price of a European option which is an essential derivative. Derivatives themselves are defined as a financial instrument whose value depends on the values of other, more basic, underlying variables.³ Regarding the Black-Scholes-Model, the volatility of the underlying asset is, beside other inputs, the only variable that can't be directly observed on the market. Consequently, one needs to forecast the volatility of the underlying asset from the date the option starts until the date the option expires in order to fully calculate the price of the option.

Since the volatility and forecasted volatility are viewed as vital model-inputs, due attention was given to find models which describe the behaviour of the volatility itself. Numerous scholarly articles have been written, describing different approaches to estimate the volatility and to take care of its characteristics, such as volatility clustering and leverage effect. These features and the corresponding models will be discussed in detail later in this paper. Furthermore, we will take a unique approach in our analysis, focusing on the so called implied volatility. The goal is to determine if the implied volatility contains any additional information about the future realized volatility compared to model based forecasts (MBFs).

¹ Brooks, 2008, p. 383.

² Poon & Granger, 2003, p. 478.

³ Hull, 2009, p. 1.

1.2 Earlier research

One of the first to investigate whether GARCH models can be used for volatility forecasting purposes was Taylor (1986), who was followed by Akgiray (1989). Akgiray (1989) came to the conclusion that GARCH constantly performs better than EWMA and models based on historical volatility derived from standard deviation of past returns over a fixed interval. Until now there has been copious research made in the field of forecasting volatility, and the general outcome has not been unanimous. Some studies find no definite results, while other have shown opposite results. Lamoureux and Lastrapes (1993) conclude that models based on implied volatility from stock options perform better than MBFs, when trying to predict the volatility of individual stocks. Contrarily Canina and Figlewski (1993) argue that implied volatility has little explanatory power of the realized volatility on S&P 100 index options. Poon and Granger (2003) summarize and compare 93 earlier studies done in the field of volatility forecasting. The result of this comparison is that implied volatility is the method, which the majority of studies find to be most accurate for forecasting volatility.

In the past, researchers have tried to use other models based on different approaches to find forecasting methods that perform better. These methods include historical volatility models and, as mentioned above, GARCH type models. Poon (2005) argues that the number of lag volatility terms included and the weights assigned to them can be used to divide historical volatility models into two different categories; the single state and the regime switching models. The first category includes, among other models, moving averages and autoregressive models.⁴ An example of this category is the Autoregressive Moving Average Model (ARMA), which was used by Pong et al. (2004) and showed positive qualities in forecasting. They concluded that the ARMA(2,1) performs well when used with realized volatility on the foreign exchange market.

The second group contains a number of different GARCH type models. Researchers have yet to reach a final conclusion as to which of these models is optimal for forecasting volatility. Some of the more popular models in this group are the GJR-GARCH first introduced by Glosten et al. (1993), and the Exponential GARCH (EGARCH) proposed by Nelson (1991). Brailsford and Faff (1996) argue that the GJR-GARCH model, compared to the EGARCH

⁴ Hull, 2009, p 4f.

model, is the better and generally the most accurate GARCH model, while Franses and Van Dijk (1996) conclude that forecasting by GJR-GARCH cannot be recommended. Finally Taylor (2001) claims that the GJR-GARCH model outdoes a standard GARCH(1,1) model.

Another topic of discussion that has not been as widely debated during the years is whether the implied volatility potentially contains information about the expectation that the affiliated parties may have about the future. This question was later investigated by Becker et al. (2007), who used the S&P 500 implied volatility index (VIX) to determine whether the expected volatility implied by option prices contains any incremental information for forecasting volatility. They concluded that the VIX does in fact contain information incremental to that contained in individual MBFs, but not when the volatility index is compared to a set of MBFs.

Furthermore, Poon and Granger (2003) conclude that forecasting of financial market volatility is possible and that the discussion should move on to focus more on what time period can be forecasted successfully and if changes in volatility can be predicted.

1.3 Problem description

This paper will generally focus on the Swiss market, because it is one of the most unique stock markets in Europe and is especially interesting due to the fact that banks represent a bigger part of the Swiss market than in other countries. The Swiss Market Index (SMI) and the Volatility Index of the SMI (VSMI), both provided by SIX Swiss Exchange (SIX), are at the centre of our interest. Since VSMI data is only available from January 1999, we have a maximal given time period of approximately 11 years.

Based on the work of Poon and Granger (2003) we will investigate if and how well volatility can be forecasted for the SMI. By reviewing 93 different papers about volatility forecasting, they concluded that most of the papers favour the implied volatility as the most favourable method to forecast volatility. However, they also argue that this outcome should not be generalized. This is one reason why we will examine the compatibility of these different approaches on the SMI volatility.

Moreover, we investigate if the VSMI provides any information beyond that which is captured in the model-based forecasts. The working paper of Becker et al. (2007) is used as a

reference to see whether the SMI options markets are capable of expecting movements in volatility unanticipated by MBFs.

1.4 Purpose

Our hypothesis is that the VSMI dominates all other MBFs and furthermore contains more information than what is captured by the used models. This means that the main purpose of the thesis is to investigate and determine the adequacy of using the VSMI for forecasting future volatility on the SMI. One part of this study aims to investigate whether the VSMI is the best method for forecasting future volatility on the SMI. The second function of the paper is to examine whether implied volatility contains additional information when compared to the set of MBFs.

1.5 Problem delimitation

The only major delimitation we face is the method used to compute the realized volatility. As we will describe later, intraday data is needed in order to avoid biases in the realized volatility. Due to the fact that this data is not easily accessible, we were forced to use daily low and high data of the SMI to compute the daily realized volatility.

1.6 Disposition

The structure of the paper will be as follows: the necessary theoretical background will be presented in chapter two, giving the reader the opportunity to fully understand the theories that are applied to our empirical observations. We will start by explaining different types of volatility and continue with their characteristics, before we introduce the models and measures that will be used further along in the thesis. In chapter three, we will describe what tests we will use and how they are executed, followed by a motivation and discussion of the methodology. Chapter four will explain what data will be used and what characteristics the raw data contains. We will also discuss the indices used and explain how these are calculated. In chapter five we present the results of the performed tests as well as discuss the interpretation of these results. Finally, in chapter six, we draw conclusions from our results and discuss what further research can be conducted in this particular field of finance.

2 Theoretical background

This section of the thesis will give a very basic overview of the fundamental theory. We try to give an insight into the different volatility types and how they can be forecasted. Additionally, we discuss ways of evaluating the quality of a forecast and how to look for evidence of additional information in the implied volatility in relation to the realized volatility.

In our thesis we generally use continuously compounded returns as shown in equation (1) instead of arithmetic returns. On the one hand this has the simple advantage that the returns are capable of being totalled and on the other hand longer time series are, in theory, more likely to be normally distributed.⁵ Equation (2) shows the mean of the compounded returns.

$$r_t = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln S_t - \ln S_{t-1} \quad (1)$$

$$\bar{r} = \frac{\sum r_t}{T} = \frac{\sum(\ln S_t - \ln S_{t-1})}{T} = \frac{\ln S_T - \ln S_0}{T} \quad (2)$$

2.1 Types of volatility

2.1.1 Historical volatility

As stated before, in finance, the volatility often refers to the variance, σ^2 , or standard deviation, σ , of returns. The standard deviation is simply the square root of the variance. Equation (3) describes the variance computed from a set of observations where \bar{r} represents the mean return.

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{t=1}^N (r_t - \bar{r})^2 \quad (3)$$

2.1.2 Realized volatility

Figlewsky (2004) states that the sample average return, \bar{r} , is a very noisy estimate of the true parameter μ . He argues that this could lead to a bias in the estimation of the volatility. As equation (2) shows, the mean is not dependent on the available observations. Only the first

⁵ Steiner & Bruns, 2002, p. 54f.

and the last observations have effect on the accuracy of the estimation. Other researchers, such as Black (1976), recommend to just assume a value for the mean rather than trying to estimate the mean from the data. Perry (1982) calculated the monthly volatility simply by summing the squared daily returns, which implies an assumed mean of zero.

Additionally, the general availability of high-frequency observations allowed researchers to find an advanced way to measure volatility. Finally, Andersen et al. (2001) and Barndorff-Nielsen and Shepard (2004) developed an estimator for ex post realized variance called realized volatility. Generally, the realized volatility is estimated by summing the squared intraday high-frequency returns. In theory, the realized volatility should be free from measurement error as the sampling frequency of the returns approach infinity. Andersen et al. (2001) argue that in practice the realized volatility should be estimated by using a 5-minute return horizon. This should moderate the contamination of market microstructure frictions, including price discreteness, infrequent trading and the bid-ask bounce.

Nevertheless, since intraday data is not easily accessible and also expensive, we calculate the daily realized volatility by high and low values of the daily stock price. Equation (4) shows how the realized volatility is computed for the SMI.

$$RV_t = \ln\left(\frac{SMI_{high,t}}{SMI_{low,t}}\right) \quad (4)$$

2.1.3 Implied volatility

As stated above, volatility is the only input variable to the Black-Scholes pricing model, that can't be observed at the market. Equations (5) to (8) provide the mathematical framework of the famous pricing model. The variable c represents the European call price, p is the European put price, S_0 is the stock price at time zero, K is the strike price, r continuously compounded risk-free rate, σ is the stock price volatility and T is the time to maturity of the option. The function $N(x)$ is the cumulative probability distribution function for a standardized normal distribution.

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2) \quad (5)$$

$$p = Ke^{-rT} N(-d_2) - S_0 N(-d_1) \quad (6)$$

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \quad (7)$$

$$d_2 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \quad (8)$$

Instead of using the volatility as input variable, traders often use the, in options implied, volatility to infer the volatility of the underlying stock. The implied volatility is therefore used to monitor the market's opinion about the volatility of a certain stock. Yet there is no method to derive the volatility from the Black-Scholes pricing model and to express the implied volatility as a function.⁶

2.2 Characteristics of volatility

Volatility has many different characteristics that may have an impact when using observations to forecast future volatility. Diebold et al. (1998) discuss the impact of such factors as current level of volatility, how long of a period you need to forecast and the intensity of volatility. Other factors that may affect the outcome of the forecasting models is the volatility clustering, leverage effect and the relationship between volatility and strike price, also known as the volatility smile. This will be further discussed below.

2.2.1 Volatility clustering

Mandelbrot (1963) was the first to conclude that “large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes”. This pattern of varying volatility over time can be seen in most financial returns series. Connolly and Stivers (2005) investigated this and found that two possible reasons for volatility clustering are new information being available on the market and the level of uncertainty. These seem to be two logical explanations since equities tend to be more volatile around the time when their financial reports are released. To be able to account for this volatility characteristic, there is a need for models that take this behaviour into consideration. Such models will be presented in the coming chapters.

⁶ Hull, 2009, p. 296f.

2.2.2 Leverage effect

As Bouchaud et al. (2001) explain, the leverage effect can be described as the negative correlation that can be observed between past returns and future volatility. This is empirically shown when the conditional variance is more strongly affected by a negative shock in an antecedent period than a positive shock. They further argue that this correlation is different between individual stocks and equity indices. For stock indices the correlation is stronger than for individual stocks. However, the correlation seems to decay faster.

2.2.3 Volatility smile

According to Hull (2009) the theory of volatility smiles builds upon the ideas of the Black-Scholes model. Following the Black-Scholes model, put-call parity is used to show the relationship between put and call prices of an option. The put-call parity between two options with an underlying asset q , including dividend yield, on the asset is

$$p + S_0 e^{-qT} = c + K e^{-rT} \quad (9)$$

Assuming that, p_{BS} and c_{BS} are the values of European put and call options that holds for the Black-Scholes model at a particular level of the volatility. This naturally means that for the two options we have the following relationship

$$p_{BS} + S_0 e^{-qT} = c_{BS} + K e^{-rT} \quad (10)$$

Put-call parity assumes that arbitrage is not possible. This means that put-call parity for two options can be translated into options which hold for the market, so that

$$p_{mkt} + S_0 e^{-qT} = c_{mkt} + K e^{-rT} \quad (11)$$

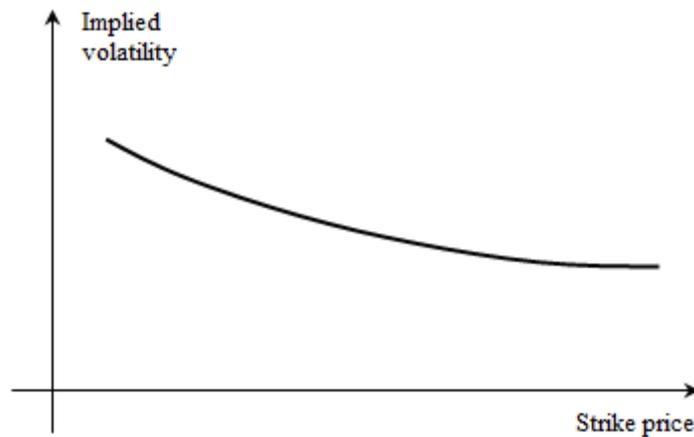
By combining equation (10) and (11), we get

$$p_{BS} - p_{mkt} = c_{BS} - c_{mkt} \quad (12)$$

Supposing that the Black-Scholes model holds, this means that the pricing error, when pricing a European put option, should be exactly the same as the pricing error, when used to price a European call option with the same time to maturity and strike price. The volatility used to price the options should therefore be the same when strike price and maturity is applied together with a given implied volatility.

The conclusion that can be drawn from this is that the relationship between strike price and time to maturity, also known as the volatility smile, is the same for calls and puts and that the volatility term structure also must be the same for the two. Rubinstein (1985) was the first to study the volatility smile in equities, concluding that the volatility decreases as the strike price increases. This relationship for equities can be seen in figure 1 below.

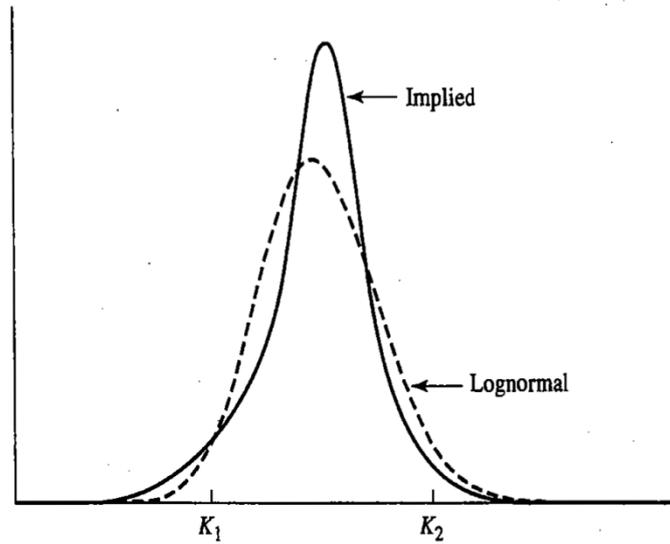
Figure 1 - Volatility smile for equities



2.2.4 Distribution characteristics

Theodossiou (2000) has, among others, concluded that distributions of log-returns of financial assets such as equities show unique characteristics such as skewness and leptokurtosis. Skewness implies that one of the distribution's tails will be longer than the other. Leptokurtosis, on the other hand, means that the distribution of log-returns of a stock has fat tails and exhibits a thinner but higher peak at the mean. According to Hull (2007) the volatility smile for equity options follows an implied probability distribution. Figure 2 shows that this distribution exhibits the very same leptokurtosis and skewness as for other financial assets. The implied distribution has a heavier left tail than the right one and has an excess peakness at the mean.

Figure 2 - Implied distribution and lognormal distribution for equity options



2.3 Modelling volatility

2.3.1 Naïve models

One of the most basic models one can use for forecasting the future volatility is simply to take the historical realized volatility and apply as a forecast. This means that the model does not take into consideration the time varying characteristics of volatility. The naïve model used in this thesis builds upon the calculated realized volatility for the last 30 days. A further discussion of the calculation of realized volatility is found in 2.1.2. Applying this on a 30 day period gives the average daily volatility equation as:

$$\hat{\sigma}_T = \frac{\sum RV_t}{T} \quad (13)$$

where $T = 30$ in this paper.

2.3.2 Autoregressive Moving Average (ARMA)

A model that lets the current variable y depend on only the past values of the same variable y plus an error term is called an autoregressive (AR) model. An $AR(p)$ model with p lags, can be expressed as:

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + u_t \quad (14)$$

An even simpler model is the Moving Average (MA) model, where the dependent variable y only depends on current and past values of a white noise disturbance term. A white noise process is defined in (15) and an MA(q), where q is the number of lags, is denoted in equation (16).

$$E(u_t) = \mu$$

$$\text{var}(u_t) = \sigma_u^2 \quad (15)$$

$$u_{t-r} = \begin{cases} \sigma_u^2 & \text{if } t = r \\ 0 & \text{otherwise} \end{cases}$$

$$y_t = \mu + \sum_{i=1}^q \theta_i u_{t-i} + u_t \quad (16)$$

As Brooks (2008) explains, an Autoregressive Moving Average (ARMA) model is constructed by combining the models in equation (14) and (16). This model uses both the previous values of the dependent variable as well as a combination of current and previous values of a white noise error term. The ARMA model demonstrates characteristics both from the autoregressive and moving average parts, showing both a geometrically decaying autocorrelation function and geometrically decaying partial autocorrelation function. The AR(p) process becomes dominant only after q lags, since this is when the autocorrelation function stops depending on the MA(q). An ARMA model can be explained by the following formulation:

$$\phi(L)y_t = \mu + \theta(L)u_t \quad (17)$$

where

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (18)$$

and

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q \quad (19)$$

where the mean of an ARMA process is given by:

$$E(y_t) = \frac{\mu}{1 - \phi_1 - \phi_2 - \dots - \phi_p} \quad (20)$$

Choosing the correct number of lag variables can be a rather time-consuming task; therefore we choose to use the model argued by Pong et al. (2004). They argue that the ARMA model most suitable for forecasting purposes is the ARMA(2,1), since the two autoregressive lags can confine the longer lasting periods of volatility.

A problem that often appears in ARMA-type models is non-stationarity. This means that there is at least one unit root present in the autoregressive part of the autoregressive process.⁷ To test for this, Dickey and Fuller (1979) developed a test to decide whether a unit root is present or not. In brief, the test is performed by estimating the following regression:

$$y_t = \phi y_{t-1} + u_t \quad (21)$$

The null hypothesis that $\phi = 1$, and the one-sided alternative that $\phi < 1$. The test can then be augmented to include p lags of the dependent variable.⁸

2.3.3 Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model is related to ARMA. The difference, denoted by the I in the acronym, is that the characteristic equation has a root on the unit circle. This condition is also known as an integrated autoregressive process. An ARMA(p,q) model in which the variable is differenced d number of times is equivalent to an ARIMA(p,d,q) model on the original data.⁹ As Enders (2004) explains, the first difference of any ARIMA(p,I,q) series has the stationary infinite-order moving average representation:

$$y_{t+s} = y_t + \alpha_0 s + \sum_{i=1}^s e_{t+i} \quad (22)$$

when

$$e_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \beta_3 \varepsilon_{t-3} + \dots \quad (23)$$

⁷ Campbell, Lo & MacKinlay, 1997, p. 64-65.

⁸ Brooks, 2008, p. 327-329.

⁹ Brooks, 2008, p. 233.

2.3.4 Autoregressive Conditional Heteroscedasticity (ARCH)

The assumption that the variance of errors is constant might not be applicable in the field of finance. Since errors in financial time series tend to vary over time, researchers felt for a long time that there was a need for a model which has an adaptable error variance. Engle (1982) was the first to present such a model, calling it an Autoregressive Conditionally Heteroscedasticity (ARCH) model. This article was the beginning of a whole new era of models.

As mentioned above, it is important to understand the time varying random variable u_t and its' conditional variance. Brooks (2008) argues that to fully understand the ARCH model, one should first divide the definition into conditional and unconditional variance of the random variable. When assumed that $E(u_t) = 0$, the conditional variance of u_t is:

$$\sigma_t^2 = \text{var}(u_t | u_{t-1}, u_{t-2}, \dots) = E[u_t^2 | u_{t-1}, u_{t-2}] \quad (24)$$

This means that the conditional variance of a zero mean normally distributed random variable u_t is equal to the conditional expected value of u_t . When this conditional variance is used in a full-scale ARCH model, the second part of the model is used to estimate the conditional mean. This equation basically describes how the dependent variable y_t varies over time. With both parts of the models stated next to each other, the complete ARCH model is:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + u_t \quad (25)$$

where $u_t \sim N(0, \sigma_t^2)$ and with the conditional variance expressed as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2 \quad (26)$$

This is known as an ARCH(q) model.

2.3.5 Generalized ARCH (GARCH)

It did not take long until researchers found severe flaws in ARCH-type models. One big difficulty that the GARCH-family of models solves is the problem of overfitting. It has always been one of the biggest challenges when estimating ARCH models to decide how many lags of the squared residual term to include. Bollerslev (1986) and Taylor (1986) found a solution when they presented the GARCH model. Since this model was first presented, researchers have developed the GARCH model further with vital additions and alterations to the original formula.

As Campbell et al. (1997) explain, the GARCH model allows the conditional variance to be dependent on its' previous own lags. The conditional variance equation of a simple GARCH(1,1) model is provided in equation (22):

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (27)$$

This model can now be extended to a GARCH(p,q), where p is the number of lags of the conditional variance and q is the number of lags of the squared error. This gives the following formulation of the conditional variance:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (28)$$

When put into the full scale model, a GARCH(p,q) has the following appearance:

$$R_t = \mu + u_t \quad (29)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (30)$$

$$u_t \sim N(0, \sigma_t^2) \quad (31)$$

One of the drawbacks of this GARCH(p,q) model is the possible problem of negative volatility. This means that the estimated volatility in a GARCH(p,q) model can be smaller than zero, which is meaningless. To solve this problem, one possible method is to impose non-negativity constraints so that all model coefficients are required to be non-negative. Another drawback is the lack of possibility to account for leverage effects which are discussed in section 2.2.2.¹⁰

2.3.6 Exponential GARCH (EGARCH)

Nelson (1991) presents a model that is known as the Exponential GARCH (EGARCH). This model has many advantages in comparison to the original GARCH model. Most importantly, the conditional variance σ_t^2 will be positive regardless if the parameters are positive or negative.¹¹ This is due to the modelling of $\log(\sigma_t^2)$. EGARCH makes it thus unnecessary to impose non-negativity constraints on the model parameters. A second important advantage is that the EGARCH can adapt to asymmetries. If the relationship between volatility and returns is nega-

¹⁰ Brooks, 2008, p. 404.

¹¹ Campbell, Lo and MacKinlay, 1997, p. 487-488.

tive, γ , will be negative.¹² The conditional variance equation of an EGARCH can be described as:

$$\ln(\sigma_t^2) = \omega + \beta * \ln(\sigma_{t-1}^2) + \gamma \frac{e_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[\frac{e_{t-1}}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right] \quad (32)$$

where the volatility clustering effect is captured by the parameter α , while the leverage effect is obtained by γ . Henceforth, the asymmetrical effect is not quadratic but exponential, which means that if $\gamma < 0$ then there is a leverage effect.¹³

2.4 Quality evaluation of forecasts

2.4.1 Root Mean Square Error (RMSE)

To be able to determine which model actually is the best at forecasting volatility, we will use different quality evaluation measures, one of them being Root Mean Square Error (RMSE). The RMSE builds upon the Mean Squared Error (MSE) measure, and is simply the root of this model. As the MSE, the RMSE aggregates the residuals into one single measure of predictive power.¹⁴ The measure can be explained as:

$$RMSE = \sqrt{\frac{1}{T - (T_1 - 1)} \sum_{t=T_1}^T (y_{t+s} - f_{t,s})^2} \quad (33)$$

where T is total sample size, including both in-sample and out-of-sample observations. T_1 is the first out-of-sample forecast observation and $f_{t,s}$ is an s -step ahead forecast of a variable made at time t .

2.4.2 Mean Absolute Error (MAE)

Another measure of the quality of forecast is the Mean Absolute Error. It measures the average absolute forecast error. As with the other quality measures, it is important to use the out-

¹² Brooks, 2008, p. 406.

¹³ For a further discussion regarding the EGARCH, see for example Nelson (1991) or Liu and Morley (2009).

¹⁴ Brooks, 2008, p. 252.

comes of the measures for comparison between different forecasting models, and not to the values of the measures individually. MAE is depicted as:

$$MAE = \frac{1}{T - (T_1 - 1)} \sum_{t=T_1}^T |y_{t+s} - f_{t,s}| \quad (34)$$

2.4.3 Mean Absolute Percentage Error (MAPE)

Building on the MAE measure, Mean Absolute Percentage Error is the third quality measure we will use in this thesis. Recommended by Makridakis (1993), the MAPE has advantages compared to the MSE such as that the outcome can be interpreted as a percentage error. The MAPE equation is defined as:

$$MAPE = \frac{100}{T - (T_1 - 1)} \sum_{t=T_1}^T \left| \frac{y_{t+s} - f_{t,s}}{y_{t+s}} \right| \quad (35)$$

2.4.4 Adjusted Mean Absolute Percentage Error (AMAPE)

Both MAE and MSE can experience problems with asymmetry between the actual and the forecasted values. This problem is solved by the AMAPE measure.¹⁵ AMAPE is expressed as:

$$AMAPE = \frac{100}{T - (T_1 - 1)} \sum_{t=T_1}^T \left| \frac{y_{t+s} - f_{t,s}}{y_{t+s} + f_{t,s}} \right| \quad (36)$$

2.4.5 Theil-U statistic

Theil (1966) presented the less commonly used *U*-statistic. The *U*-statistic shows how the model in question performs in comparison to the benchmark model. A value of one indicates that the considered model and the benchmark model perform equally well. If the value is below one, this means that the model is more accurate than the benchmark. The Theil-U metric can be expressed as:

¹⁵ Brooks, 2008, p. 253.

$$U = \frac{\sqrt{\sum_{t=T_1}^T \left(\frac{y_{t+s} - f_{t,s}}{y_{t+s}} \right)^2}}{\sqrt{\sum_{t=T_1}^T \left(\frac{y_{t+s} - fb_{t,s}}{y_{t+s}} \right)^2}} \quad (37)$$

Where $fb_{t,s}$ is the forecast obtained from a benchmark model.¹⁶

2.4.6 LINEX

Discussed by Varian (1975), Zellner (1986) and Christoffersen and Diebold (1997) for example, the LINEX loss function is given by:

$$LINEX = \frac{1}{I} \sum_{i=1}^I \left[\exp\left(-\alpha(\hat{\sigma}_i^2 - \sigma_i^2)\right) + \alpha(\hat{\sigma}_i^2 - \sigma_i^2) - 1 \right] \quad (38)$$

where α is a given parameter. The unique aspect of the LINEX loss function is that positive errors are weighted differently from the negative errors as Yu (2002) explains. If $\alpha > 0$, this mean that the LINEX loss function is about linear for $\hat{\sigma}_t^2 - \sigma_t^2 > 0$ and exponential for $\hat{\sigma}_t^2 - \sigma_t^2 < 0$. It is thus an asymmetric quality measure that needs to be taken into account when applied.

3 Methodology

In this section, we provide an explanation for how the questions will be investigated. We start by addressing the issues of matching time horizons and how to forecast realized volatility. The discussion moves on to how to investigate the forecast quality and determine if the implied volatility contains any additional information in relation to MBFs. Finally, we present a discussion regarding the choice of method and some criticism of it. To investigate the discussed problems, we use EViews 6. The source code of the used program is given in the appendix.

¹⁶ Brooks, 2008, p. 255.

3.1 Matching time horizons

As partly shown in chapter 2.3, there are a number of different models that can be used to analyse and forecast volatility. Before we define which models we use to model the realized volatility of the SMI we first want to introduce a basic rule which we follow.

As described above, the VSMI bases its volatility calculations on a 30-day basis but is expressed as a yearly volatility, which will be explained closely in chapter 4.3. This means that, if we want to compare this volatility index to MBFs, some adjustments have to be made.

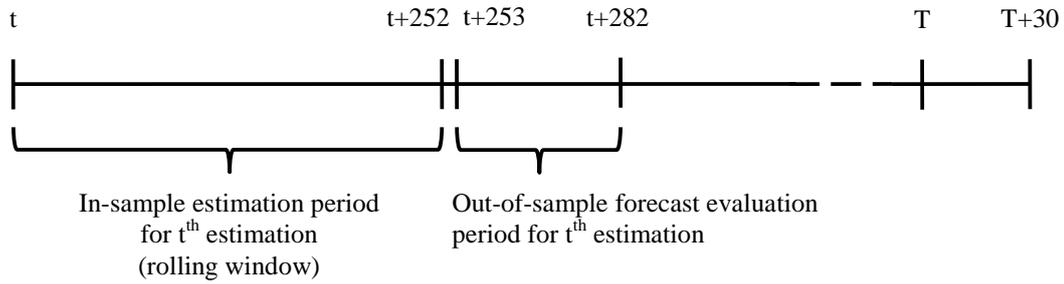
To make the data comparable, we first transform the VSMI into daily volatility. Second, we create MBFs for each day, for 30 days ahead, in order to calculate the average daily volatility over that time period. And, finally, we introduce a 30-day average value for the computed realized volatility. These adjustments allow us to compare the forecasts and the realized volatility with the volatility index.

3.2 Forecasting realized volatility

After calculating the realized volatility by using the natural logarithm of daily high and low prices, as shown in the theoretical part, we now want to go further into the details on how we estimate our future realized volatility.

The fact that we use historical data spanning from January 4th 1999 to April 1st 2010 for our calculations, leads us to the question of the window sizes for our estimation. One would agree that it is not useful to estimate a model over a multiple-year time period to then forecast just a few days. We therefore decided to estimate the models over a time period of 252 observations, which represents a year. These models are, as suggested by Brooks (2008), used to forecast an out-of-sample period of 30 days. To cover the whole data range, a rolling window approach will be used for the total sample period, meaning that the start and end date of the in-sample is flexible and successively moves as new 30-day forecasts are made. The general case for the forecasts is shown in figure 3 below, where $1 \leq t \leq T$ and $T = 2802$. The first 252 observations are used as the first in-sample-period for the first estimation and the last 30 observations are used as the out-of-sample period for the last estimation, which results in a total number of 2551 estimations per model.

Figure 3 - Estimation periods



Following Pong et al. (2004), the first model we use to forecast the realized volatility is the simple ARMA(2,1) model. Since the contained MA(1) part of the model dies out after a one-day forecast, a multiple-day-ahead forecast would lead to a decreasing estimated volatility. To work around this drawback we simply forecast one day at a time, assemble the forecasted volatility into the rolling window, drop the last observation and re-estimate the model. In that way it is possible to generate 30 forecasts. The reader should be aware that the original data is not altered and after the 30 forecasts are transformed to the average 30-day-ahead value, the rolling window is placed back to then move just by one observation. Another drawback is that the ARMA(2,1) model could forecast negative realized volatilities. Again, we try to compensate this problem and set a negative forecast to zero.

In Chapter 5.1 one can see that the Augmented Dickey-Fuller (ADF) test suggests unit roots in 201 estimation attempts. Therefore we follow Brooks (2008) and try to overcome this problem by forecasting the realized volatility with an ARIMA(1,1,1) model even though Chu (1978) does not find the model very accurate for longer forecasts. Again, to work around the problem of the dying MA(1) process, we use the same approach as described above and re-estimate the model with another rolling-window to create more accurate forecasts of the realized volatility.

The last model we use to estimate the realized volatility is of the family of the Autoregressive Conditional Heteroscedasticity (ARCH) models. Since the Generalized ARCH (GARCH) model would probably give negative values for the variance, we therefore decided to use the Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model.

3.3 Investigating the forecast quality

To be able to conclude the quality of the forecast, we will use theoretical analysis methods such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Adjusted Mean Absolute Percentage Error (AMAPE). Furthermore, more advanced, less commonly used performance measures, namely the Theil-U statistic and the LINEX loss function to investigate the quality of our forecasts. Since the Theil-U statistic computes the performance of a forecast in relation to a reference, we use a very simple Naïve model as a benchmark. This Naïve model computes the current value simply by the average of the last 30-day observations.

3.4 Predictive power and information content

It is important to see how well the forecasts actually correspond to the actual future observations to be able to fully determine the sufficiency of a model. One way of doing this is to estimate a regression where the independent variable is the forecasted value, and the dependent variable is the realized volatility for the date which is forecasted. To evaluate the accuracy of the forecast, one can look at the R^2 and compare it to the outcomes of other models. This method can be used as one way of determining which model is the best at forecasting volatility. A simple regression fulfilling this purpose is as follows:

$$\overline{RV}_{t+30} = \alpha + \beta \hat{\sigma}_t + u_t \quad (39)$$

where \overline{RV}_{t+30} is the average realized volatility over 30 days, $\hat{\sigma}_t$ is the average 30-day forecast at time t and u_t is the error term.

When performing this method for investigating the best forecast model, there is always a risk of facing correlated residuals and/or heteroscedasticity. For example, Christensen and Prabhala (1998), Fleming et al. (1995) and Fleming (1998) all experience these problems of biased estimators. The Ordinary Least Squares (OLS) method that is used to obtain the value of the regression is only an efficient estimator if these problems are non-existent. A solution to this problem is to correct the standard errors as proposed by Newey & West (1987).

To test whether the forecasting models contain useful information, one can adjust the simple regression used above, resulting in:

$$RV_{t+1} = \alpha + \beta \hat{\sigma}_t + \varepsilon_t \quad (40)$$

where $\hat{\sigma}_t$ is the forecasted volatility at time t and RV_{t+1} is the realized volatility one day ahead. The idea behind the regression is to determine whether the forecasted volatility today says anything about the realized volatility of tomorrow. It is constructed first and foremost for this purpose and not for determining the accuracy of the forecasts. We will perform this test on all forecasting models and then rank the models based on the outcome.

The regression is estimated by using OLS as well as the very same method as used above presented by Newey & West (1987) in order to correct the standard errors. The result is then depicted by the slope coefficient β , and if $\beta > 0$ there is a positive relationship between the forecasted volatility of today and the realized volatility of tomorrow.

3.5 Additional information in VSMI

In the last part of this paper, we will investigate whether the VSMI contains information incremental to that captured by MBFs. Following the approach of Becker et al. (2007), we divide the VSMI into two components. The first component is a parameter containing the same information as can be found in the combination of MBFs called $VSMI_t^{MBF}$. The other component is a parameter containing any other information that can be useful for forecasting purposes called $VSMI_t^*$. This means that the VSMI can be expressed as:

$$VSMI_t = VSMI_t^{MBF} + VSMI_t^* \quad (41)$$

This is made to ensure orthogonality between $VSMI_t^*$ and ω_t . One can use a linear projection to map $VSMI_t$ into ω_t with $VSMI_t^{MBF} = QVSMI_t$, where Q is the projection matrix of the stacked volatility forecasts as well as the current realized volatility of the SMI in ω_t . Becker et al. (2007) further explain that $VSMI_t^*$ is given by $VSMI_t - VSMI_t^{MBF}$, which by definition is orthogonal to the elements in ω_t . This can also be regarded in practical terms as a linear regression:

$$VSMI_t = \gamma_0 + \boldsymbol{\gamma}_1 \boldsymbol{\omega}_t + \varepsilon_t \quad (42)$$

with

$$VSMI_t^{MBF} = \hat{\gamma}_0 + \hat{\boldsymbol{\gamma}}_1 \boldsymbol{\omega}_t \quad (43)$$

and

$$VSMI_t^* = \hat{\varepsilon}_t \quad (44)$$

If the volatility index encompassed all of the information in the MBFs one would expect $VSMI_t^*$ to not contain any incremental information and to be orthogonal to future realized volatility. In that case $VSMI_t^{MBF}$ forecasts future volatility equally well as $VSMI_t$. The authors argue that a number of strategies can be used to implement the orthogonality test between the future realized volatility and $VSMI_t$. Since Harvey and Newbold (2000) argue that an OLS regression of $VSMI_t^*$ as the independent and future realized volatilities as an independent variable followed by a F-test of the null hypothesis that the coefficient of the future realized volatilities is zero, is just applicable for short-term forecasts, we decide to directly use a GMM optimization, which is used by Becker et al. (2007) to verify their results.

A GMM estimate of $\boldsymbol{\gamma} = (\gamma_0, \boldsymbol{\gamma}_1)'$ in equation (42) minimizes $V = \mathbf{M}'\mathbf{H}\mathbf{M}$, where $\mathbf{M} = T^{-1}(\varepsilon_t(\boldsymbol{\gamma})'\mathbf{Z}_t)$ is a $k * 1$ vector of moment conditions, \mathbf{H} is a $k * k$ weighting matrix and \mathbf{Z}_t is a vector of instruments. One can now choose H to be the variance-covariance matrix of the k moment condition in \mathbf{M} to be able to minimize coefficient variances, this according to Hamilton (1994), among others.

To see if the residuals in (42) are uncorrelated with elements in \mathbf{Z}_t , one can perform the test for overidentifying restrictions $J = T\mathbf{M}'\mathbf{H}\mathbf{M}$, with a distribution of $\chi^2(k - \dim(\boldsymbol{\gamma}))$, where k is the number of instruments and $\dim(\boldsymbol{\gamma})$ is the number of estimators. Whenever $k > \dim(\boldsymbol{\gamma})$, the null hypothesis cannot be rejected and the residuals are thus uncorrelated with the elements in \mathbf{Z}_t . \mathbf{Z}_t was constructed to include the MBFs in $\boldsymbol{\omega}_t$, in combination with the information regarding future realizations of volatility \mathbf{RV}_{t+30} , $\mathbf{Z}_t = (\boldsymbol{\omega}_t', \mathbf{RV}_{t+30}')'$.

The future realizations of volatility in \mathbf{RV}_{t+30} capture information regarding actual realization of volatility during the 30 trading days following time t . It is also reasonable to include RV over shorter horizons to test if the VSMI holds incremental information in respect to shorter horizons. So we define \mathbf{RV}_{t+30} as follows:

$$\mathbf{RV}_{t+30} = \{\overline{RV}_{t+1}, \overline{RV}_{t+5}, \overline{RV}_{t+10}, \overline{RV}_{t+15}, \overline{RV}_{t+20}, \overline{RV}_{t+25}, \overline{RV}_{t+30}\} \quad (45)$$

where \overline{RV}_{t+j} is the average realized volatility between $t+1$ and $t+j$. Furthermore, Becker et. al (2007) suggest the investigation of additional information within the volatility index regarding

the changes of volatility. Therefore we define another set of instruments, shown in the equation below.

$$dRV_{t+30} = \{d\overline{RV}_{t+1}, d\overline{RV}_{t+5}, d\overline{RV}_{t+10}, d\overline{RV}_{t+15}, d\overline{RV}_{t+20}, d\overline{RV}_{t+25}, d\overline{RV}_{t+30}\} \quad (46)$$

where $d\overline{RV}_{t+j}$ is the average realized volatility between $t+1$ and $t+j$ minus the realized volatility at time t .

The J test was used to test the null hypothesis that $VSMI_t^*$ is orthogonal to all elements in \mathbf{Z}_t , which is equivalent to that the $VSMI_t$ is not containing any incremental information relevant to the future realizations of volatility.

3.6 Choice of method

Poon and Granger (2003) compared different studies of volatility models to investigate which models generally gave better forecasts. The outcome of their research paper is not completely unanimous, which is why this paper uses different estimators to see which one can explain the data in the best way. However, the article states a number of different models to be significantly more popular in the literature than others, which is why we have chosen to use these models for analysing and forecasting the data. To be able to determine if there is relevant information in the forecasted volatility, we will follow the approach of Becker et al. (2007), which examines the S&P 500 implied volatility index (VIX). They conclude that the implied volatility contains no additional information compared to a set of MBFs. Therefore, our intention is to investigate if this is true for the Swiss stock market.

3.7 Criticism of the method

There are some obvious problems when choosing this methodology for forecasting volatility. The choice of time period will undoubtedly affect the results, as Diebold et al. (1998) discuss. A crisis that has been described as the worst crisis since the Great Depression¹⁷ will affect the results and the chosen forecast period making the results more difficult to interpret than if a stable period was used.

¹⁷ see WSJ 2008-09-08 “Worst Crisis Since ‘30s, With No End Yet in Sight”.

The method might also be more efficient if the forecasting period was divided into a “normal” period with lower levels of volatility and an “exceptional” period with more volatile observations. This method would enable the models to be used on two different scenarios and thus a deeper analysis could be made.

Finally, more econometric models could be used for forecasting and examining the data to further support our results.

4 Data description

This chapter will provide an insight on the data used in this paper. SMI and the VSMI will also be briefly introduced to give the reader a deeper understanding of the background and conditions of the Swiss Market.

4.1 General data description

The observation period starts on January 4th 1999 and ends on April 1st 2010, totalling more than 11 years. The data for the SMI and VSMI was both provided by the Thomson Reuters Datastream database. We obtain daily high and low prices for the SMI to be able to calculate the realized volatility and daily close prices for the VSMI. After removing holidays and other non-trading days, the total number of daily observations sums up to 2832.

The reason for starting in 1999 is simply that data from the VSMI is only available this far back. However, the number of observations should still be sufficient to perform the intended hypothesis tests. The observations include more fluctuant times such as the dot-com bubble, as well as the strong positive market trend in the middle of the decade. Both the SMI and the VSMI indices are provided by SIX Swiss Exchange Ltd (hereafter SIX).

4.2 Swiss Market Index (SMI)

The SMI is the most important stock index in Switzerland. It represents the 20 largest equities and covers around 85 % of the total capitalization of the Swiss equity market. The composition of the index is reviewed once a year, meaning that firms that have lost their position as

one of the 20 largest equities will be swapped against another equity. The annual rankings are based on both turnover and capitalization.¹⁸

Calculation of the index is made in real time, so that every new transaction that is made relating a stock on the SMI will induce the index to change. The index was standardized on June 30th 1988 at 1 500 points, and is calculated by only taking the tradable portion of the shares into consideration. It is also capital-weighted, so that the largest firms affect the index more than the relatively small firms.¹⁹

4.3 VSMI as a volatility index

First introduced on April 20th 2005, the Volatility Index on the SMI (VSMI) makes it possible for investors to trade pure volatility on the SMI. For statistical purposes, the VSMI is calculated back to January 4th 1999. The index is calculated by using a replicating portfolio that responds directly to volatility changes.

The VSMI uses the very same estimation methods as other volatility indices such as the VSTOXX and the VDAX-NEW. It utilizes implicit variances to all Eurex-traded SMI options of the same durations and is regulated on the basis of a constant residual term of 30 days.²⁰ The maximum time of the options to expiration is two years, while the VSMI is calculated on the basis of eight expiry months.

Besides the main index VSMI, SIX also provides sub-indices for each time to expiration of the SMI options ranging from one month up to two years, which are calculated and distributed for all models. All options available are used when calculating these sub-indices, and are based on the best ask and best bid on each of these options in the Eurex system.²¹

To obtain the main index, one uses linear interpolation of the two sub-indices that together enclose a remaining period of 30 days to expiration for the VSMI. This means that the main

¹⁸ SIX Swiss Exchange Ltd., 2010a, p. 4.

¹⁹ SIX Swiss Exchange Ltd., 2010a, p. 4.

²⁰ SIX Swiss Exchange Ltd., 2010b, p. 1.

²¹ SIX Swiss Exchange Ltd., 2009, p. 1.

index does not expire and that the estimation method helps to decrease effects that typically result in strong volatility fluctuation the closer the options come to expiration.²²

Each sub-index is estimated until two days prior to expiration. SIX disseminates each new sub-index for the first time on the second day of trading of the relevant SMI options. This results in the calculation of the VSMI main index to be expressed as:²³

$$\begin{aligned}
 VSMI &= 100 * \sqrt{\left[T_i * \sigma_i^2 \left[\frac{N_{T_{i+1}} - N_T}{N_{T_{i+1}} - N_T} \right] + T_{i+1} * \sigma_{i+1}^2 * \left[\frac{N_T - N_{T_i}}{N_{T_{i+1}} - N_{T_i}} \right] \right] * \frac{N_{365}}{N_T}} \\
 &= \sqrt{\left[T_i * VSMI_i^2 * \left[\frac{N_{T_{i+1}} - N_T}{N_{T_{i+1}} - N_T} \right] + T_{i+1} * VSMI_{i+1}^2 * \left[\frac{N_T - N_{T_i}}{N_{T_{i+1}} - N_{T_i}} \right] \right] * \frac{N_{365}}{N_T}}
 \end{aligned} \tag{47}$$

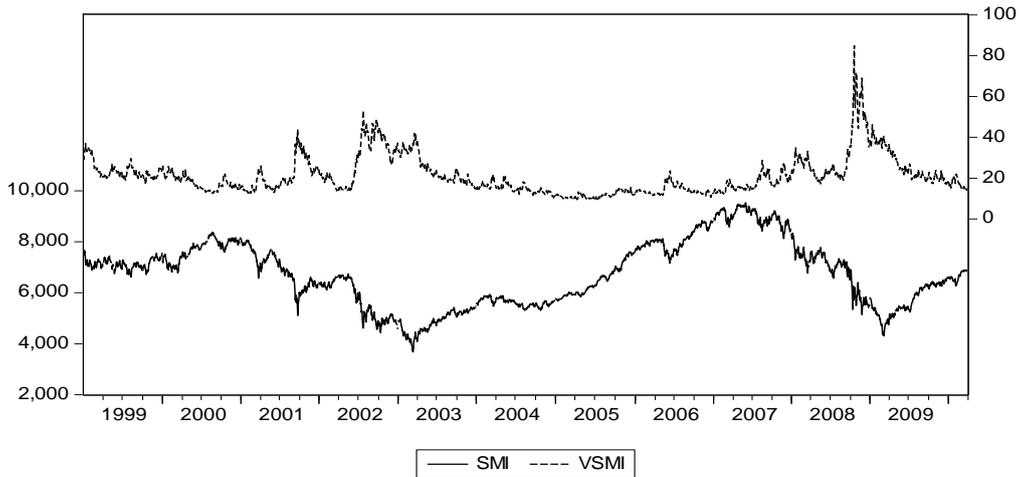
where N_{T_i} is the time to expiration of the i^{th} best bid and best ask of all SMI-options (OSMI), $N_{T_{i+1}}$ is the time to expiration of the $(i+1)^{\text{th}}$ OSMI, N_T is the time for next x days and N_{365} is the time for a standard year.

When comparing the historical developments of the SMI and the VSMI, one can see that the indices show a highly negative correlation. The two indices are compared in figure 4 where the SMI is depicted in given points while the VSMI is in percentages.

²² SIX Swiss Exchange Ltd., 2009, p. 2.

²³ SIX Swiss Exchange Ltd., 2009, p. 8.

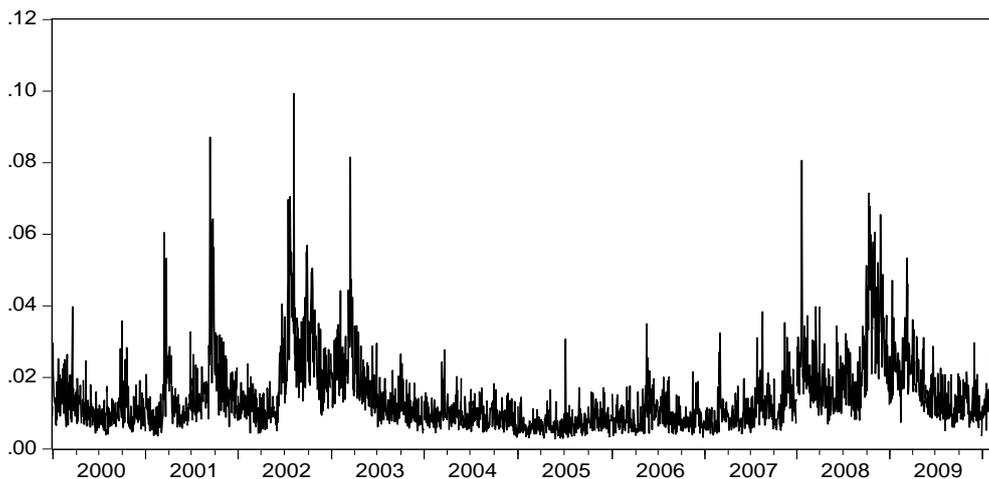
Figure 4 - SMI and VSMI



4.4 Descriptive Statistics

To verify that the data does not contain too many errors, we conduct a graphic inspection. To further see the characteristics of this particular set of observations, the realized volatility is calculated for a better overview. Figure 5 provides a time series view of the daily realized volatility of the SMI. The 2832 observations show both, periods of high volatility, such as the time period from the middle of 2002 to the middle of 2003 or from the end of 2007 to the middle of 2009, as well as periods of lower volatility, such as in 2004 and 2005.

Figure 5 - SMI daily realized volatility – time series view



This distribution of the realized volatility shows the typical characteristics of a financial times series as discussed in 2.2.4. The observations have a mean of the realized volatility of 1,43%. A normal distribution has a kurtosis of 3, while the kurtosis of the distribution in figure 6 has

a value at almost 13 implying fat-tails.²⁴ The skewness of the SMI daily realized volatility is around 2,5, while a normal distribution should have a skewness of 0.²⁵

Figure 6 - SMI daily realized volatility - descriptive statistics

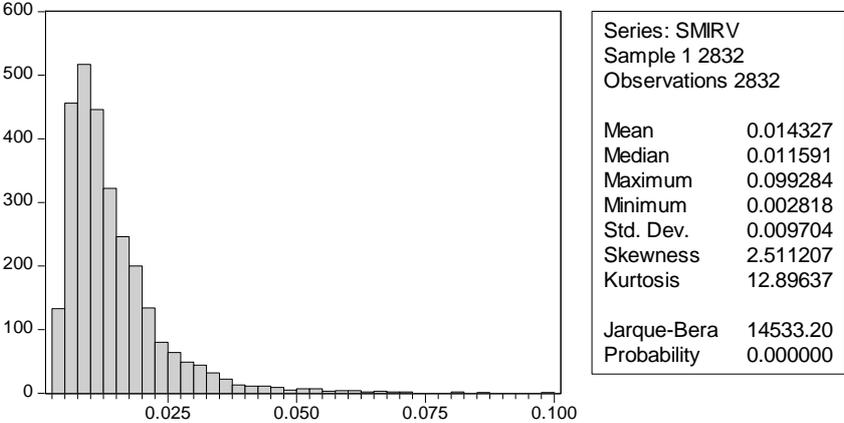
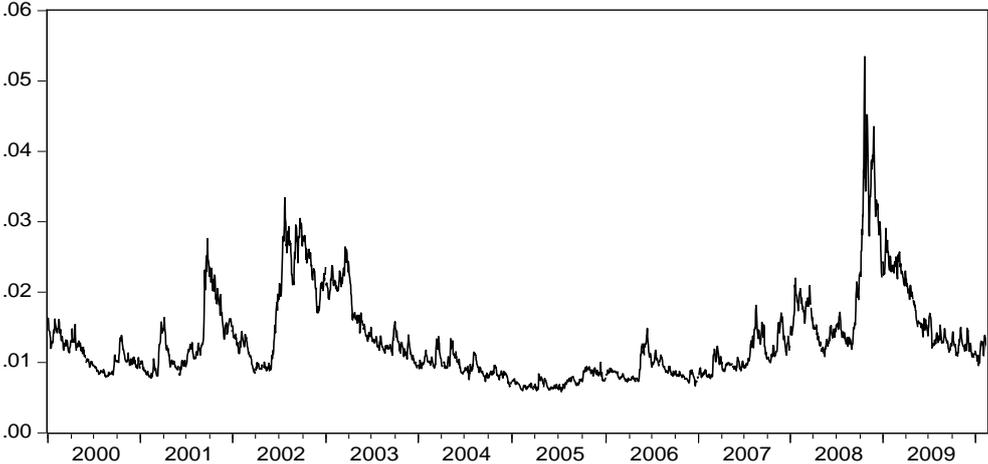


Figure 7 shows the 30-day ahead average realized volatility of the SMI. When comparing this figure to figure 5, it is clear that this figure shows a lower volatility in general. The highest peak of the 30-day average is at the end of 2008, while the daily realized volatility shows its' highest value during the fall of 2002.

Figure 7 - SMI 30-day realized volatility average - time series view

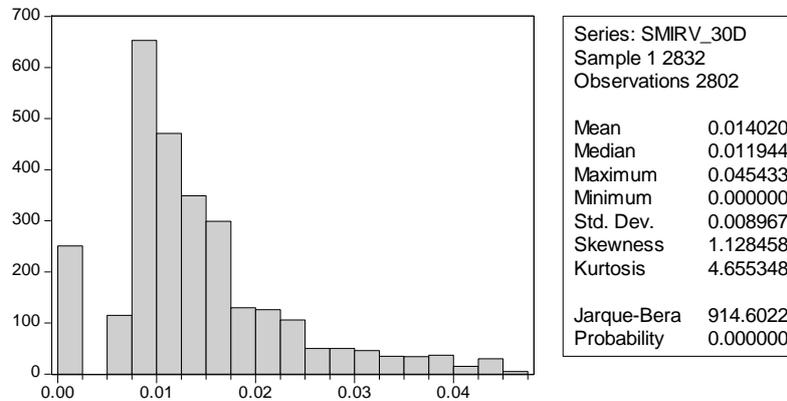


The distribution of the 30-day calculated realized volatility average also shows the characteristics of a financial time series as in figure 8. However, both the skewness and kurtosis are lower, while the mean is approximately the same as the mean of the daily realized volatility

²⁴ Campbell, Lo & MacKinlay, 1997, p. 16.
²⁵ Brooks, 2002, p. 161.

average. The standard deviations of the two distributions are similar, with a value of about 0,009 for both the daily realized volatility and the 30-day realized volatility.

Figure 8 - SMI 30-day realized volatility average - descriptive statistics



5 Empirical results

5.1 Descriptive statistics of rolling estimation period

The descriptive statistics in this chapter differ from the ones provided in Chapter 4.4, in both length and placement of the observation window. The data characteristics above were computed for an observation window, reaching from December 29th 1999 to February 18th 2010, while the following descriptive statistics apply on the rolling window of 252 daily observations. Moreover, the characteristics for the rolling window were calculated 2551 times to get a better understanding of the behaviour of the SMI during a one-year period.

Table 1 below describes the realized volatility of the SMI. One can see that the realized volatility ranges from 0,282% to 9,93% and that the mean realized volatility varies from 0,43% to 5,70%. Since we know that the volatility on financial markets is not constant over time, this outcome was expected. If we take a look at the distribution parameters of the realized volatility, it gets clear that the distribution is leptokurtic and positively skewed. Also the Jarque-Bera test indicates that the assumption of normality in this distribution needs to be rejected. Brooks (2008) supports this finding and states that non-normality is a common property of financial data, even though various financial models and theories ignore that fact.

Table 1 - Realized volatility - Average descriptive statistics*

	Mean	Stddev	Min	Max	Kurt	Skew	JB	Prob
Mean	1,44%	0,76%	0,43%	5,70%	9,66	1,93	882,2	0,00%
Stddev	0,54%	0,36%	0,12%	2,57%	4,79	0,66	1020,0	0,00%
Min	0,71%	0,31%	0,28%	1,82%	3,05	0,78	25,6	0,00%
Max	2,63%	1,48%	0,94%	9,93%	29,50	4,03	8057,4	0,00%

*Sample size is 252, estimation done 2551 times

In table 2 we provide the result of the ADF test for non-stationarity. In 201 cases the null hypothesis of the observation period containing a unit root could not be rejected. If we compare this number to the total amount of 2551 estimations, we see that the ADF-test indicates a unit root for 7,88% of the estimations. Even though we expected the number of found unit roots to be under 128 (5%) we can still argue that the null of a unit root can be rejected on a 10% significance base.

Table 2 - Realized Volatility - Augmented Dickey-Fuller tests*

	Lag Length	T-Stat	Probability
Mean	1,97	-6,17	1%
Min	0,00	-15,53	0%
Max	6,00	-1,70	43%
Stddev	1,45	3,36	3%
Detected unit roots			201

*Sample size is 252, estimation done 2551 times

5.2 Model estimation

To create MBFs we first estimate three different models to capture the behaviour of the realized volatility. Recall that the rolling estimation period is 252 days and is moved 2551 times. The first model we estimate is the ARMA(2,1) model. It is directly applied on the daily realized volatility of the SMI. Since the forecasting method, to create 30 forecasts for the daily realized volatility, imply a re-estimation of the model, the data below represents the average model variables of the 2551 initial estimations. As we can see in table 3, the coefficients show a wide range of values. This is due to the fact that the estimation periods contain unit roots. Especially the constant C is taking on extreme values which deviate from the mean and median by far. On the other hand, C was just included in the estimation to ensure that the OLS estimators are unbiased and plays therefore no role in forecasting itself. The MA values were estimated by backcasting. Looking at the overall significance of the coefficients, one can see that the AR(2) process is just significant in 22% of the estimations.

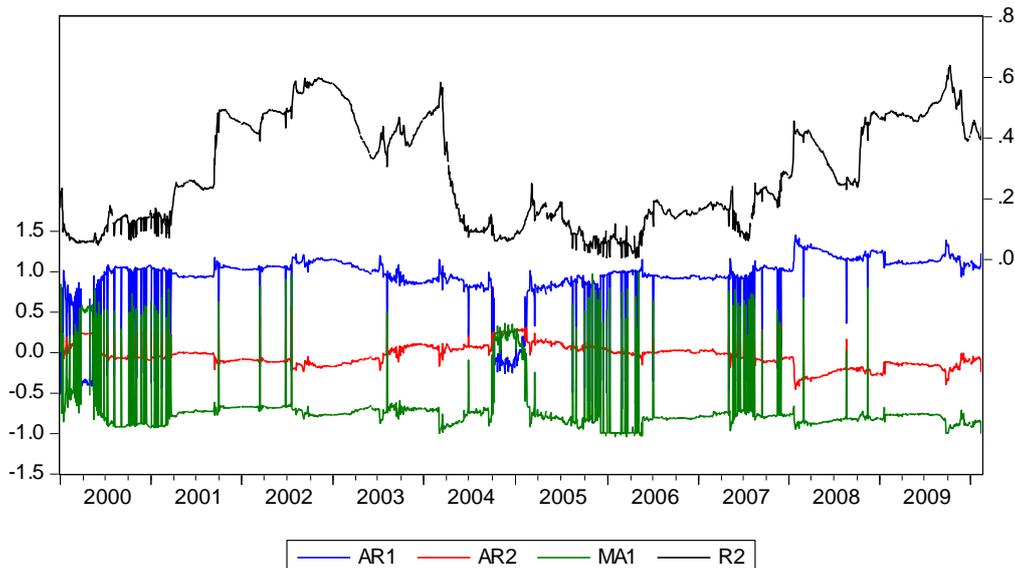
Table 3 - Estimated ARMA(2,1) parameters*

	C	T-Stat	Prob	AR1	T-Stat	Prob	AR2	T-Stat	Prob	MA1	T-Stat	Prob
Mean	0,05	10,24	1,7%	0,89	9,11	4,2%	-0,03	-0,27	34,9%	-0,67	-58,01	2,8%
Median	0,01	7,98	0,0%	1,00	9,38	0,0%	-0,04	-0,56	29,2%	-0,77	-9,59	0,0%
Stddev	7,33	7,85	11,5%	0,40	4,36	15,9%	0,14	1,82	31,0%	0,37	2123,84	11,9%
Min	-254,31	0,00	0,0%	-0,91	-14,44	0,0%	-0,45	-6,80	0,0%	-1,04	-107251,99	0,0%
Max	264,02	49,54	100,0%	1,45	23,91	99,5%	0,57	9,55	99,9%	0,99	235,35	99,9%
Signif			97,1%			91,1%			22,0%			92,5%

*parameters represent initial estimation. Sample size is 252, estimation done 2551 times

In figure 9 we plot the collected starting parameters of the ARMA(2,1) model. The constant C is left out since it was not used to create forecasts. Generally, we can observe that the parameters are relatively constant. However, one can see from the very low R^2 values and value-switching coefficients of the AR(1) and MA(1) processes, that the model fails to have a constant fit on multiple time-periods. This is probably due to the fact that the coefficients of the AR(1) and MA(1) are both close to 1 and -1 respectively. The coefficient for of AR(2) process on the other hand seems to be very stable and does not take on high values at all.

Figure 9 - Initial ARMA(2,1) parameters



While the ARMA(2,1) seems to have stability issues, the ARIMA(1,1,1) model seems to be able to provide more stable coefficients. The constant C shows a value of zero through all estimations and is with an overall significance of 6% not relevant. Also the AR(1) process is just in 27% of the estimations significant. The most dominant coefficient is the one of the MA(1) process which is significant in all estimations and always takes on a negative value. The T-statistics of the MA(1) coefficient take on huge values.

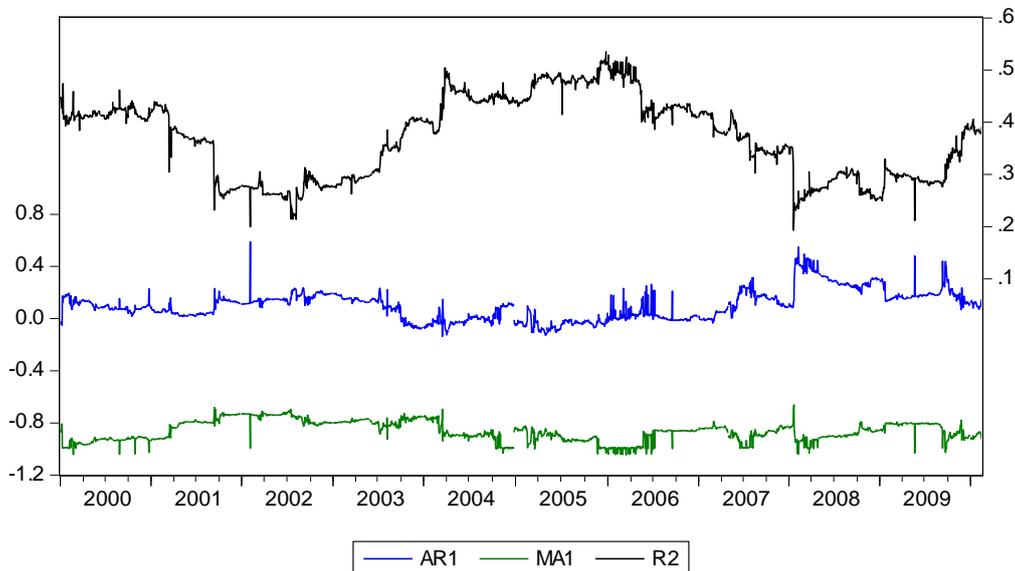
Table 4 - Estimated ARIMA(1,1,1) parameters*

	C	T-Stat	Prob	AR1	T-Stat	Prob	MA1	T-Stat	Prob
Mean	0,00	-0,02	60%	0,10	1,30	32%	-0,86	-162,45	0,00%
Median	0,00	-0,01	66%	0,10	1,27	19%	-0,86	-22,92	0,00%
Stddev	0,00	1,12	29%	0,11	1,60	32%	0,08	6413,44	0,00%
Min	0,00	-11,94	0%	-0,14	-1,79	0%	-1,04	-323955,02	0,00%
Max	0,00	6,58	100%	0,59	12,72	100%	-0,66	-5,76	0,00%
Signif			6%			27%			100%

*parameters represent initial estimation. Sample size is 252, estimation done 2551 times

Figure 10 shows the relatively constant coefficients of the ARIMA(1,1,1) model. Again, we can observe big changes in the R^2 over time.

Figure 10 - Initial ARIMA(1,1,1) parameters



The last model we estimate is the EGARCH model. Generally, the coefficient alpha, which covers the symmetric reaction to shocks in the past, seems to be least significant with 31,7%. The model is able to capture the asymmetry of shocks since the coefficient Gamma is negative and significant most of the time. A drawback is that the model seems to have trouble to fit. The t-values take on high values again.

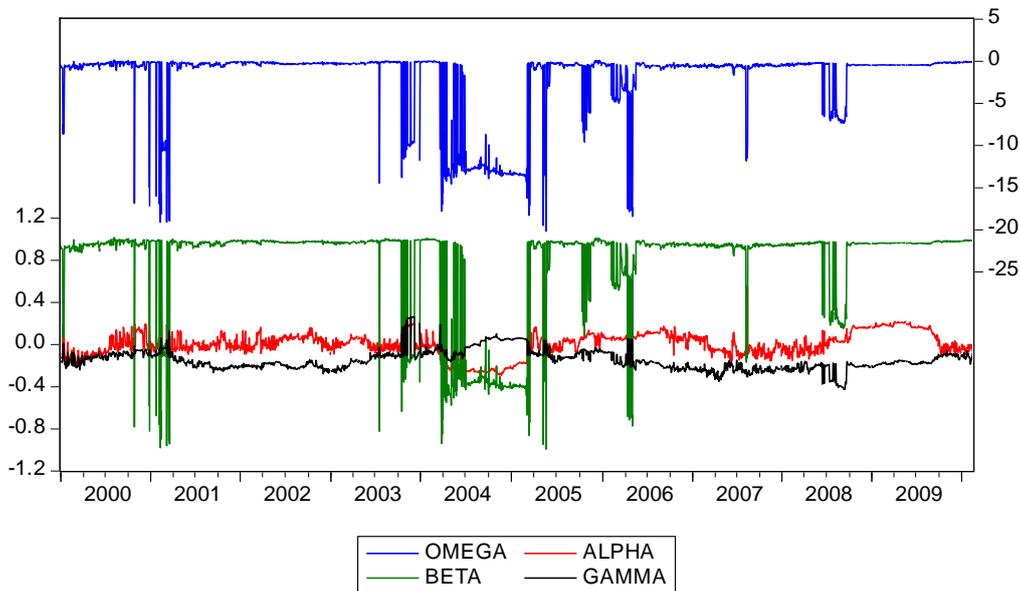
Table 5 - Estimated EGARCH parameters*

	Omega	T-Stat	Prob	Alpha	T-Stat	Prob
Mean	-2,06	-1211988,70	7,6%	0,00	-2797266,19	28,9%
Median	-0,41	-2,69	0,7%	0,00	0,04	15,4%
Stddev	4,19	26727724,24	15,9%	0,11	68293921,39	30,4%
Min	-20,14	-1244918335,20	0,0%	-0,29	-3169163812,28	0,0%
Max	0,17	20220331,11	100,0%	0,55	7,60	100,0%
Signif			73,7%			31,7%
	Beta	T-Stat	Prob	Gamma	T-Stat	Prob
Mean	0,78	118315827,57	6,2%	-0,15	-123077,80	7,7%
Median	0,96	62,10	0,0%	-0,17	-3,93	0,0%
Stddev	0,44	2964479343,35	17,3%	0,10	3694358,45	19,7%
Min	-0,99	-87,33	0,0%	-0,43	-142979598,78	0,0%
Max	1,02	112916547462,86	99,6%	0,27	7,50	99,9%
Signif			86,6%			80,8%

*parameters represent initial estimation. Sample size is 252, estimation done 2551 times

The problem mentioned above is visible in figure 11. One can see that the normally stable coefficients show big drops which do not seem to be connected to special events in the financial markets like the financial crisis that started in the end of 2007.

Figure 11 - Initial EGARCH parameters



5.3 Descriptive statistics of generated forecasts

This part of the paper discusses the descriptive statistics of the MBFs, the naïve model used for benchmarking as well as the VSMI and the actual realized volatility of the SMI. Table 6 gives a general view on the descriptive statistics on the average daily 30-day realized volatili-

ty forecast, which were computed 2551 times. The null hypothesis of a normal distribution can be rejected for all the datasets. Also all the series show a leptokurtic and positively skewed distribution. As discussed before, this is a common characteristic of financial data. It turns out that the EGARCH model produces very low forecasts for the future realized volatility. Among the MBFs, the ARIMA(1,1,1) model seems to have the highest correlation with the average realized volatility.

Table 6 – Average 30-day forecast results*

	SMIRV	VSMI	ARMA	ARIMA	EGARCH	NAIVE
Mean	1,54%	1,31%	1,16%	1,43%	0,71%	1,54%
Median	1,27%	1,14%	0,96%	1,11%	0,67%	1,28%
Maximum	4,54%	5,35%	5,92%	6,07%	2,38%	4,54%
Minimum	0,57%	0,58%	0,05%	0,50%	0,11%	0,57%
Std. Dev,	0,82%	0,60%	0,68%	0,86%	0,30%	0,82%
Skewness	1,61	1,82	1,72	2,04	1,29	1,61
Kurtosis	5,25	7,36	7,66	7,82	5,76	5,25
Corr(SMIRV)	1,00	0,75	0,67	0,74	0,60	0,69
Corr(VSMI)	0,75	1,00	0,87	0,91	0,71	0,92
Jarque-Bera	1640,2	3431,3	9083,1	4240,2	1523,4	1639,3
Probability	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

*30-day forecasts, 2551 times estimated

The graph view as well as the distribution can be seen in the appendix figure 1 and appendix figure 2.

5.4 Forecast evaluation

In this section we provide the results of the forecast quality evaluation. We compare the computed forecasts with the 30-day average realized volatility as well as the standard deviation of the returns of the SMI. The results in table 7 show the value for each used quality measure and the rank the model scored is given in brackets. The models are compared to the 30-day average realized volatility of the SMI as well as the standard deviation of the SMI returns.

Table 7 - Forecasting Quality

Quality compared to RV of upcoming 30 days						
	ARMA	ARIMA	EGARCH	VSMI	NAIVE	
MSE	0,000053 (4)	0,000038 (2)	0,000116 (5)	0,000035 (1)	0,000043 (3)	
MAE	0,011375 (5)	0,003913 (2)	0,008500 (4)	0,003357 (1)	0,004273 (3)	
AMAPE	17,63 (4)	11,87 (2)	35,28 (5)	10,19 (1)	12,12 (3)	
MAPE	55,45 (5)	22,18 (2)	50,50 (4)	17,65 (1)	24,62 (3)	
Theil-U	1,19 (4)	0,79 (2)	2,80 (5)	0,54 (1)	1,00 (3)	

Quality compared to SD of upcoming 30 days						
	ARMA	ARIMA	EGARCH	VSMI	NAIVE	
MSE	0,000035 (2)	0,000048 (3)	0,000048 (4)	0,000029 (1)	0,000062 (5)	
MAE	0,004002 (2)	0,004734 (3)	0,004866 (4)	0,003852 (1)	0,005797 (5)	
AMAPE	16,76 (3)	16,73 (2)	23,77 (5)	15,32 (1)	20,43 (4)	
MAPE	33,75 (1)	41,65 (4)	38,10 (3)	35,75 (2)	54,41 (5)	
Theil-U	0,38 (1)	0,62 (4)	0,41 (3)	0,41 (2)	1,00 (5)	

Generally, we conclude that the VSMI dominates all other models when being compared to the realized volatility of the SMI and the standard deviation. Only the MAPE and Theil-U measures indicates that the VSMI is the second best model after the ARMA(2,1) model, when compared to the standard deviation. Besides this finding, the results seem to be somewhat inconsistent. If we focus on the more advanced Theil-U statistic, which measures the performance of the models in relation to the naïve model as a benchmark model, we can see that the VSMI model and the ARIMA(1,1,1) model can outperform the naïve model for explaining the realized volatility. The ARMA(1,1) model almost performs as well as the naïve model and the EGARCH model produces a far worse statistic than the benchmark model. When the forecasts of the different models are being compared to the standard deviation, the picture looks different. All models perform better than the simple naïve model. The VSMI is just producing the second best statistic after the ARMA(2,1) model.

Let us now turn to the LINEX loss function, which weights positive errors differently from negative errors. If the given parameter $\alpha > 0$, the function is approximately linear for over-predictions and exponential for under-predictions. Thus, negative errors receive more weight, which leads to the behaviour that under-predictions are taken into account more seriously. This is consistent with the basic understanding of the volatility theory. Brailsford and Faff (1996) argue that the under-prediction of volatility, which corresponds to the call option price, is a greater concern to a seller than a buyer. Based on Hwang et al. (1999) we assumed the values of the parameter α to be -20, -10, 10 and 20. Table 8 presents the value and ranking under the four LINEX loss function. It is clear that LINEX delivers very similar results for different assumptions of the parameter α . When the models are compared to the realized vola-

tility of the SMI, LINEX ranks the VSMI and ARMA(2,1) best, while EGARCH falls to the last place. Just when $\alpha = 20$, and therefore over-estimations are penalized, ARMA(2,1) and ARIMA(1,1,1) swap places. Interestingly, when we compare the forecasting results to the standard deviation of the SMI returns, the EGARCH model takes second place.

Table 8 - Forecasting Quality – LINEX

Quality compared to RV of upcoming 30 days (LINEX)*							
α	ARMA	ARIMA	EGARCH	VSMI	NAIVE		
-20	0,0231552 (2)	0,0234110 (3)	0,0350943 (5)	0,0175465 (1)	0,0240620 (4)		
-10	0,0058043 (2)	0,0058496 (3)	0,0088073 (5)	0,0044019 (1)	0,0060230 (4)		
10	0,0058359 (2)	0,0058437 (3)	0,0088756 (5)	0,0044328 (1)	0,0060383 (4)		
20	0,0234081 (3)	0,0233644 (2)	0,0356402 (5)	0,0177935 (1)	0,0241844 (4)		

Quality compared to SD of upcoming 30 days (LINEX)*							
α	ARMA	ARIMA	EGARCH	VSMI	NAIVE		
-20	0,0140651 (3)	0,0271841 (4)	0,0130817 (2)	0,0108763 (1)	0,0272148 (5)		
-10	0,0035165 (3)	0,0067721 (4)	0,0032820 (2)	0,0027227 (1)	0,0067919 (5)		
10	0,0035174 (3)	0,0067250 (4)	0,0033055 (2)	0,0027299 (1)	0,0067686 (5)		
20	0,0140722 (3)	0,0268076 (4)	0,0132694 (2)	0,0109347 (1)	0,0270288 (5)		

*values were multiplied by 1000

5.5 Predictive Power

The following regressions are made using OLS and the standard errors are adjusted using Newey-West, to consider heteroscedasticity and autocorrelation. The regression has the average 30-day realized volatility of the SMI as a dependent variable and the VSMI as a daily volatility, as well as average 30-day volatilities generated by the different models as independent variables. For the estimation we had 2551 observations reaching from December 29th 1999 to February 18th 2010. An ADF test for that period shows that there is no unit root in the realized volatility or 30-day average for that time period. Table 9 shows the results of the different regressions. All coefficients except the one in estimation 7 are significant. This indicates that the MBFs are biased. Regressions 1 to 4 show that all the slope coefficients for the MBFs are significant when being regressed. The large slope coefficient of regression 3 indicates again that the generated forecasts of the EGARCH model are too low. Regression 5 to 7 compare the MBFs to the VSMI. It seems that the MBFs of the ARMA(2,1) are dominated by the VSMI, since the slope coefficient of the model is not significant in regression 5. Regression 6 shows that both the ARIMA(1,1,1) MBFs and the VSMI are significant, but the coefficient is still small and indicates a bias in the forecasts. Finally, the MBFs of the EGARCH model as well as the VSMI are significant and the constant is insignificant which indicates that the two models together deliver an unbiased forecast. Generally, one can say that the

MBFs perform quite similarly as the VSMI and since all the slope coefficients are significantly different from zero, the models seem to have a predictive power regarding the future realized volatility.

Table 9 - Predictive power analysis

Regression		Constant	ARMA	ARIMA	EGARCH	VSMI	R2	Adj. R2
1	Coefficient	0,0060	0,8110				0,450	0,450
	Std Error	0,0006	0,0529					
	T-Stat	10,24	15,32					
	Prob	0,00%	0,00%					
2	Coefficient	0,0053		0,7039			0,552	0,552
	Std Error	0,0005		0,0428				
	T-Stat	9,86		16,43				
	Prob	0,00%		0,00%				
3	Coefficient	0,0037			1,6478		0,355	0,355
	Std Error	0,0010			0,1666			
	T-Stat	3,65			9,89			
	Prob	0,03%			0,00%			
4	Coefficient	0,0022				1,0128	0,557	0,557
	Std Error	0,0007				0,0579		
	T-Stat	3,09				17,49		
	Prob	0,20%				0,00%		
5	Coefficient	0,0023	0,1094			0,9061	0,559	0,559
	Std Error	0,0007	0,0595			0,0933		
	T-Stat	3,24	1,84			9,71		
	Prob	0,12%	6,61%			0,00%		
6	Coefficient	0,0031		0,3528		0,5535	0,581	0,581
	Std Error	0,0007		0,0824		0,1158		
	T-Stat	4,44		4,28		4,78		
	Prob	0,00%		0,00%		0,00%		
7	Coefficient	0,0012			0,3576	0,8878	0,565	0,565
	Std Error	0,0007			0,1460	0,0738		
	T-Stat	1,70			2,45	12,03		
	Prob	8,86%			1,44%	0,00%		

5.6 Information Content

This regression investigates whether the MBFs can explain the realized volatility one day ahead. Again, we see that the slope coefficient in regression 3 indicates that the realized volatility forecasted by the EGARCH model is too low. Furthermore, the constant is not significantly different from zero and therefore suggests that the EGARCH forecasts are unbiased estimators for the realized volatility one day ahead. The constants in regression 1, 2 and 4 are significantly different from zero and indicate point to biased estimators. Overall, we can conclude that the MBFs contain information regarding the one day ahead realized volatility of the SMI.

Table 10 - Information content analysis

Regression		Constant	ARMA	ARIMA	EGARCH	VSMI	R2	Adj. R2
1	Coefficient	0,0029	0,9912				0,453	0,453
	Std Error	0,0005	0,0484					
	T-Stat	5,37	20,46					
	Prob	0,00%	0,00%					
2	Coefficient	0,0019		0,8662			0,564	0,563
	Std Error	0,0003		0,0287				
	T-Stat	5,57		30,13				
	Prob	0,00%		0,00%				
3	Coefficient	0,0001			1,9988		0,352	0,351
	Std Error	0,0012			0,2005			
	T-Stat	0,11			9,97			
	Prob	91,27%			0,00%			
4	Coefficient	-0,0016				1,2161	0,541	0,541
	Std Error	0,0007				0,0625		
	T-Stat	-2,29				19,47		
	Prob	2,22%				0,00%		

5.7 Test for additional information in VSMI

The analysis of additional information in the VSMI on top of “general” information provided by the used models to forecast volatility is presented here. The Generalized Method of Moments (GMM) is used to test the orthogonality of the error terms of equation (41) to the realized volatility of different time horizons. The nine estimations, which include the average realized volatilities over different time horizons in the instruments, are presented in table 11. Each GMM estimation states the used instruments. The used instruments in estimation 1 are restricted to the estimators used in the regression. Therefore, the generalized method of moments optimization technique provides the same outcome as an ordinary OLS regression. This estimation therefore has no j-statistic and probability assigned. The following optimizations add more parameters to the instruments. Optimization 2 adds the vector \mathbf{RV}_{t+30} which is defined in equation (45) to the instruments while the optimizations 3 to 9 just contain one additional instrument. The j-statistic and the corresponding p-values indicate that the null hypothesis of orthogonality of $VSMI_t^*$ to the realized volatility can be rejected at a 5% significance level for most GMM optimizations. This is evidence that the VSMI actually holds more information than the used MBFs. This is somewhat inconsistent with the findings of Becker et al. (2007). The authors suggest that there is overwhelming evidence that implied volatility holds information incremental in individual MBFs, but not a wider array of MBFs is considered in the estimations. A possible explanation for our finding is that the models lack the ability to capture the characteristics of the realized volatility good enough. Another factor that

plays a role in this result is that Becker et al. (2007) were using a far wider set of MBFs in their GMM optimization and therefore covered more information within these models.

Table 11 - GMM estimation – SMIRV

Opt		C	ARMA	ARIMA	EGARCH	RV	J-Stat	Prob
1	Coefficient	0,00255	0,23293	0,45746	0,30050	-0,06057	N/A	N/A
	T-Statistic	4,78204	2,38325	3,94543	2,22040	-0,94902		
	Probability	0,00%	1,72%	0,01%	2,65%	34,27%		
	Instruments	c, ω , RV						
2	Coefficient	0,00237	0,36899	0,39474	0,16706	-0,03092	8,46	29,41%
	T-Statistic	5,42098	5,71192	5,58086	1,88620	-0,79723		
	Probability	0,00%	0,00%	0,00%	5,94%	42,54%		
	Instruments	c, ω , RV, \overline{RV}_{t+30}						
3	Coefficient	0,00242	0,35962	0,41529	0,15026	-0,03658	5,74	1,65%
	T-Statistic	5,18117	4,88915	4,19075	1,60350	-0,68987		
	Probability	0,00%	0,00%	0,00%	10,89%	49,03%		
	Instruments	c, ω , RV, \overline{RV}_{t+30}						
4	Coefficient	0,00243	0,37272	0,39568	0,14592	-0,02712	6,93	0,85%
	T-Statistic	5,34709	5,36440	4,32070	1,65762	-0,56764		
	Probability	0,00%	0,00%	0,00%	9,75%	57,03%		
	Instruments	c, ω , RV, \overline{RV}_{t+25}						
5	Coefficient	0,00246	0,37389	0,38416	0,14857	-0,02195	7,78	0,53%
	T-Statistic	5,44336	5,37257	4,39345	1,71476	-0,48695		
	Probability	0,00%	0,00%	0,00%	8,65%	62,63%		
	Instruments	c, ω , RV, \overline{RV}_{t+20}						
6	Coefficient	0,00245	0,36605	0,38734	0,16038	-0,02457	7,28	0,70%
	T-Statistic	5,40829	5,18225	4,36404	1,78513	-0,53197		
	Probability	0,00%	0,00%	0,00%	7,44%	59,48%		
	Instruments	c, ω , RV, \overline{RV}_{t+15}						
7	Coefficient	0,00257	0,35978	0,36712	0,17293	-0,01643	6,66	0,99%
	T-Statistic	5,66539	5,34125	4,53384	1,98793	-0,39071		
	Probability	0,00%	0,00%	0,00%	4,69%	69,60%		
	Instruments	c, ω , RV, \overline{RV}_{t+10}						
8	Coefficient	0,00254	0,31939	0,39696	0,21690	-0,03261	3,67	5,53%
	T-Statistic	5,38321	4,33460	4,34083	2,15714	-0,66089		
	Probability	0,00%	0,00%	0,00%	3,11%	50,87%		
	Instruments	c, ω , RV, \overline{RV}_{t+5}						
9	Coefficient	0,00244	0,24078	0,50420	0,27370	-0,08820	1,30	25,41%
	T-Statistic	4,63719	2,46105	4,34251	2,19331	-1,50548		
	Probability	0,00%	1,39%	0,00%	2,84%	13,23%		
	Instruments	c, ω , RV, \overline{RV}_{t+1}						

Significance tests were performed using the Andrews-Monahan weighting matrix with pre-whitening.

Table 12 provides the results of the GMM estimation which considers the changes in the future realized volatility as an instrument. Again, optimization 1 is the equivalent to an ordinary OLS regression. Aside the rejection of the orthogonality assumption in optimization 2 in table 12, all other optimizations cannot reject the null hypothesis of no correlation between the changes in realized volatility and the error terms. In other words there is no evidence that the VSMI contains information beyond that contained in the MBFs regarding the prediction of changes in the realized volatility.

Table 12 - GMM estimation – SMIDRV

Opt		C	ARMA	ARIMA	EGARCH	RV	J-Stat	Prob
1	Coefficient	0,00255	0,23293	0,45746	0,30050	-0,06057	N/A	N/A
	T-Statistic	4,78204	2,38325	3,94543	2,22040	-0,94902		
	Probability	0,00%	1,72%	0,01%	2,65%	34,27%		
	Instruments	c, ω , RV						
2	Coefficient	0,00202	0,11898	0,68722	0,36503	-0,18040	16,47	2,12%
	T-Statistic	3,58279	1,13375	8,13550	3,02370	-8,85848		
	Probability	0,03%	25,70%	0,00%	0,25%	0,00%		
	Instruments	c, ω , RV, dRV_{t+30} *						
3	Coefficient	0,00227	0,16322	0,67154	0,28965	-0,17763	2,23	13,52%
	T-Statistic	3,76270	1,38721	7,33645	2,01257	-5,71114		
	Probability	0,02%	16,55%	0,00%	4,43%	0,00%		
	Instruments	c, ω , RV, dRV_{t+30}						
4	Coefficient	0,00220	0,15331	0,70535	0,27564	-0,19100	2,49	11,43%
	T-Statistic	3,71369	1,30936	7,86308	1,98168	-6,35761		
	Probability	0,02%	19,05%	0,00%	4,76%	0,00%		
	Instruments	c, ω , RV, dRV_{t+25}						
5	Coefficient	0,00216	0,13976	0,73560	0,26548	-0,20177	2,82	9,30%
	T-Statistic	3,70631	1,18946	7,96873	1,99324	-6,78146		
	Probability	0,02%	23,44%	0,00%	4,63%	0,00%		
	Instruments	c, ω , RV, dRV_{t+20}						
6	Coefficient	0,00217	0,14136	0,72941	0,26894	-0,19958	2,74	9,78%
	T-Statistic	3,72028	1,21176	8,06992	2,00971	-7,02586		
	Probability	0,02%	22,57%	0,00%	4,46%	0,00%		
	Instruments	c, ω , RV, dRV_{t+15}						
7	Coefficient	0,00218	0,14501	0,71849	0,27508	-0,19621	2,65	10,36%
	T-Statistic	3,74753	1,25584	8,00366	2,04520	-7,32750		
	Probability	0,02%	20,93%	0,00%	4,09%	0,00%		
	Instruments	c, ω , RV, dRV_{t+10}						
8	Coefficient	0,00222	0,15482	0,68978	0,28533	-0,18480	2,31	12,83%
	T-Statistic	3,76223	1,34832	8,06876	2,03938	-7,02579		
	Probability	0,02%	17,77%	0,00%	4,15%	0,00%		
	Instruments	c, ω , RV, dRV_{t+5}						
9	Coefficient	0,00227	0,14820	0,68115	0,30381	-0,18451	2,46	11,67%
	T-Statistic	3,88363	1,37630	7,33554	2,28381	-7,72182		
	Probability	0,01%	16,89%	0,00%	2,25%	0,00%		
	Instruments	c, ω , RV, dRV_{t+1}						

Significance tests were performed using the Andrews-Monahan weighting matrix with pre-whitening.

6 Conclusion

The following section will provide the reader with a summary of our results and interpretations. We will summarize the findings regarding each problem described and then at the end draw more general conclusions about volatility and suggest further research.

We have used different models, such as ARMA(2,1), ARIMA(1,1,1) and EGARCH, in order to model the realized volatility. We succeeded to estimate all of the given models, even though some estimations experienced difficulties with problems due to the fact that the real-

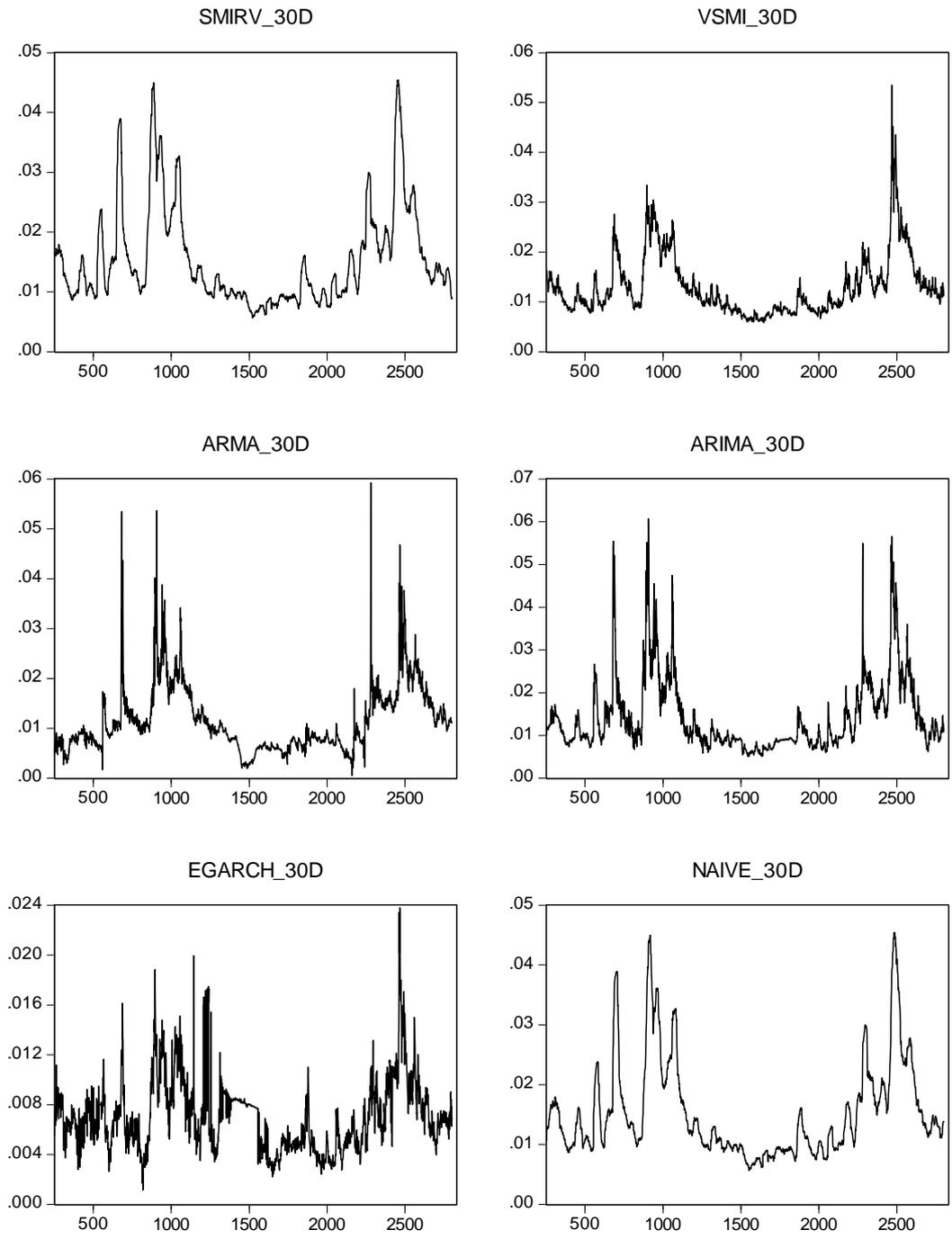
ized volatility seems to contain unit roots during some periods. The result of our comparison between the MBFs and the VSMI shows that no model was able to dominate the VSMI in terms of forecast accuracy. When comparing the models in terms of predictive power, the ARIMA model fits the data equally well as the VSMI, while the EGARCH and ARMA(2,1) model do not fit the realized volatility. Further findings of the predictive power analysis are that the VSMI significantly dominates the ARMA(2,1) model. While our findings show that the VSMI is the best predictor for realized volatility, it turns out that in our framework the VSMI is not an unbiased estimator. Furthermore, the results show that all the models succeed to demonstrate that they can provide information about the realized volatility one day ahead.

The third part of the thesis consisted of investigating whether there is additional information contained in the VSMI compared to the set of MBFs. We found evidence for incremental information stored in the VSMI. This finding is not completely in line with Becker et al. (2007). Their investigation of the additional information held within the VIX, in comparison to information contained in a wider set of MBFs, showed no excess information. One possible explanation for this outcome is that we use a smaller range of models. However, Becker et al. (2007) conclude that there is strong evidence that the implied volatility holds information incremental to that of an individual MBF. Generally, we conclude that our hypothesis holds and that the VSMI is in fact the best predictor for future volatility of the SMI.

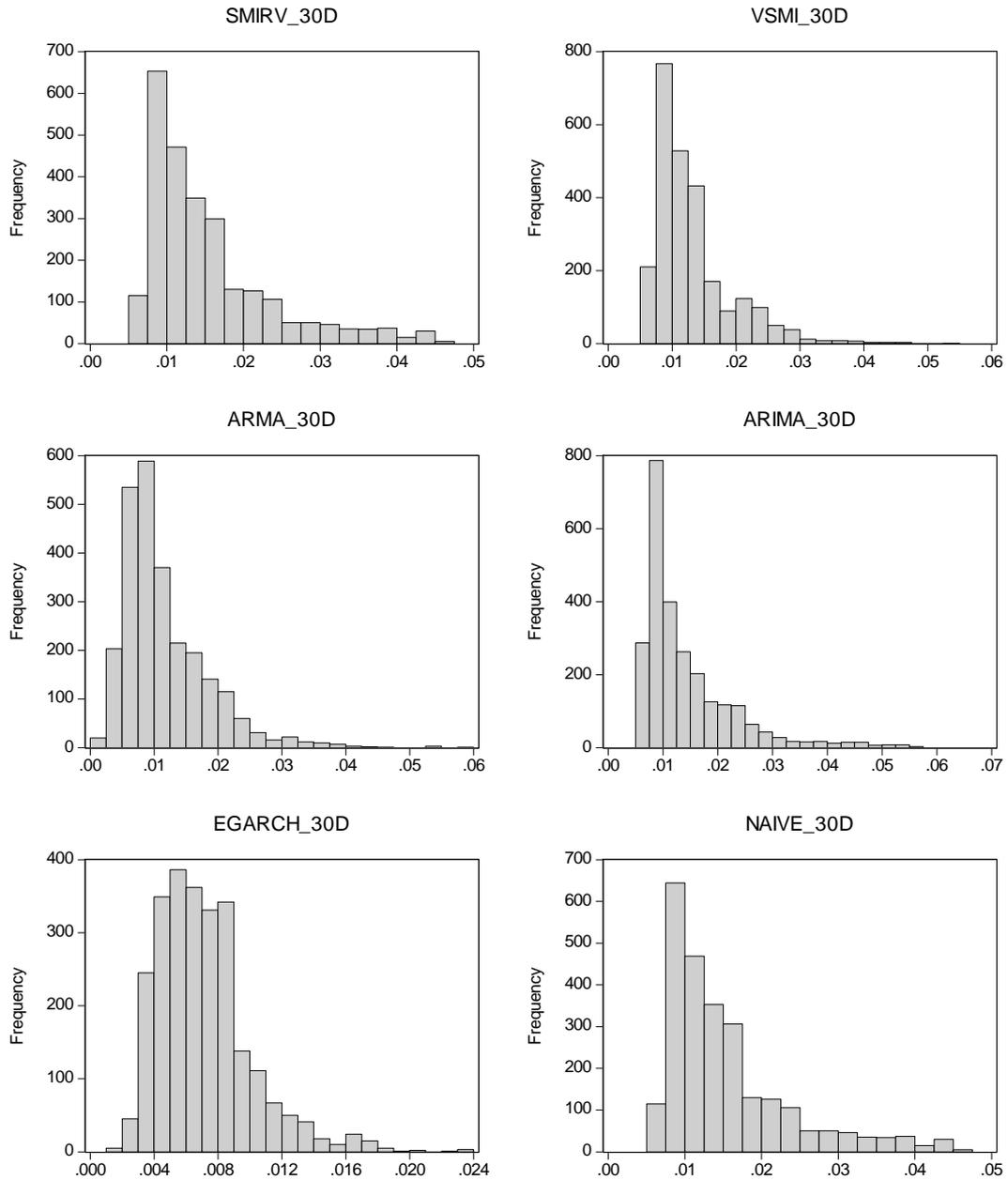
Even though this paper has provided some important insights about the VSMI and the predictability and informational contents of volatility, we made some delimitations during the process. This leaves the subject open for further research. One possible improvement could be to use intraday data to compute the realized volatility as well as to include more MBFs to increase the reliability of the results. Another way to improve the outcome would be to change the setup of time horizons to estimate the models, since our parameters seem to be unstable for a number of models. Furthermore one could consider separating the times of high volatility from times of low volatility and reviewing them separately.

7 Appendix

Appendix Figure 1 - RV, VSMI and MBFs - Graph view



Appendix Figure 2 – RV, VSMI and MBFs - Distributions



Source code for calculations – EViews 6

```

smpl @all
!observations=2832      'number of observations
!rolling_window=252    'insample observations = rolling window
!horizon=30            'forecast horizon = VSMI horizon

for %y smirv arma arima egarch vsmi naive  'variables for forecast and comparison
series(!observations-!horizon) {%y}_30d
next

series(!observations) smirv_dummy          'help variable
vector(!horizon) forecastbuffer           'buffer
matrix(!observations-!horizon,38) smirvdesstat=0 'for descriptive statistics
vector(38) smirvdesstatbuffer=0           'buffer for descriptive statistics
matrix(!observations-!horizon,8) dickey=0  'for ADF

matrix(!observations-!horizon,14) arma21parameters=0      'initial parameters ARMA(2,1)
matrix(!observations-!horizon,11) arima11parameters=0    'initial parameters ARI-
MA(1,1,1)
matrix(!observations-!horizon,14) egarchparameters=0     'initial parameters EGARCH
matrix(!observations-!horizon,!horizon) arma21forecast=0 '30 day forecasts ARMA(2,1)
matrix(!observations-!horizon,!horizon) arima11forecast=0 '30 day forecasts ARI-
MA(1,1,1)
matrix(!observations-!horizon,!horizon) egarchforecast=0 '30 day forecasts EGARCH
vector(14) arma21buffer=0
vector(11) arima11buffer=0
vector(14) egarchbuffer=0

for !i=!rolling_window to !observations-!horizon
smpl !i-!rolling_window+1 !i

'GENERATE BASIX SERIES
'=====
!rvaverage=0
!naive=0

for !j=1 to !horizon
!rvaverage=!rvaverage+smirv(!i+!j)^2 'calculate the RV for 30 days
!naive=!naive+smirv(!i-30+!j)^2
next !j

smirv_30d(!i)=@SQRT((1/!horizon)*!rvaverage)
vsmi_30d(!i)=vsmi(!i)/100*@SQRT(1/252) 'y to m and calendar day adjustment
naive_30d(!i)=@SQRT((1/!horizon)*!naive)

'CALCULATION OF DESCRIPTIVE STATISTICS
'=====
!length=252
!ljungbox_dummy=0
for !ljung=1 to 250
!ljungbox_dummy=!ljungbox_dummy+@cor(smrv,smrv(-!ljung))^2/(!length-
!ljung)
if !ljung=1 then
smirvdesstatbuffer(10)=!length*(!length+2)*!ljungbox_dummy
smirvdesstatbuffer(11)=@chisq(smrvdesstatbuffer(10),1)
endif
if !ljung=2 then
smirvdesstatbuffer(13)=!length*(!length+2)*!ljungbox_dummy
smirvdesstatbuffer(14)=@chisq(smrvdesstatbuffer(13),2)
endif
if !ljung=3 then
smirvdesstatbuffer(16)=!length*(!length+2)*!ljungbox_dummy
smirvdesstatbuffer(17)=@chisq(smrvdesstatbuffer(16),3)
endif
if !ljung=4 then
smirvdesstatbuffer(19)=!length*(!length+2)*!ljungbox_dummy
smirvdesstatbuffer(20)=@chisq(smrvdesstatbuffer(19),4)
endif
if !ljung=5 then
smirvdesstatbuffer(22)=!length*(!length+2)*!ljungbox_dummy
smirvdesstatbuffer(23)=@chisq(smrvdesstatbuffer(22),5)
endif
if !ljung=10 then
smirvdesstatbuffer(25)=!length*(!length+2)*!ljungbox_dummy
smirvdesstatbuffer(26)=@chisq(smrvdesstatbuffer(25),10)

```

```

endif
if !ljung=20 then
    smirvdesstatbuffer(28)=!length*(!length+2)*!ljungbox_dummy
    smirvdesstatbuffer(29)=@chisq(smirvdesstatbuffer(28),20)
endif
if !ljung=50 then
    smirvdesstatbuffer(31)=!length*(!length+2)*!ljungbox_dummy
    smirvdesstatbuffer(32)=@chisq(smirvdesstatbuffer(31),50)
endif
if !ljung=100 then
    smirvdesstatbuffer(34)=!length*(!length+2)*!ljungbox_dummy
    smirvdesstatbuffer(35)=@chisq(smirvdesstatbuffer(34),100)
endif
if !ljung=250 then
    smirvdesstatbuffer(37)=!length*(!length+2)*!ljungbox_dummy
    smirvdesstatbuffer(38)=@chisq(smirvdesstatbuffer(37),250)
endif
next !ljung

smirvdesstatbuffer(1)=@mean(smirv)
smirvdesstatbuffer(2)=@stdev(smirv)
smirvdesstatbuffer(3)=@min(smirv)
smirvdesstatbuffer(4)=@max(smirv)
smirvdesstatbuffer(5)=@kurt(smirv)
smirvdesstatbuffer(6)=@skew(smirv)
smirvdesstatbuffer(7)=!rolling_window/6*(@skew(smirv)^2+((@kurt(smirv)-3)^2)/4)
smirvdesstatbuffer(8)=@chisq(smirvdesstatbuffer(7),2)
smirvdesstatbuffer(9)=@cor(smirv,smirv(-1))
smirvdesstatbuffer(12)=@cor(smirv,smirv(-2))
smirvdesstatbuffer(15)=@cor(smirv,smirv(-3))
smirvdesstatbuffer(18)=@cor(smirv,smirv(-4))
smirvdesstatbuffer(21)=@cor(smirv,smirv(-5))
smirvdesstatbuffer(24)=@cor(smirv,smirv(-10))
smirvdesstatbuffer(27)=@cor(smirv,smirv(-20))
smirvdesstatbuffer(30)=@cor(smirv,smirv(-50))
smirvdesstatbuffer(33)=@cor(smirv,smirv(-100))
smirvdesstatbuffer(36)=@cor(smirv,smirv(-252))
rowplace(smirvdesstat,@transpose(smirvdesstatbuffer),!i)

smirv.uroot(adf,save=dickeybuffer)
rowplace(dickey,@transpose(dickeybuffer),!i)

'CALCULATION OF ARMA(2,1)
'=====
equation estmodel.ls smirv c ar(1) ar(2) ma(1)
!dof= estmodel.@regobs-estmodel.@ncoef 'degree of freedom

'estimate parameters for ARMA(2,1) and put them into arma21parameters(!i)
arma21buffer(1)=@coefs(1)
arma21buffer(2)=@tstats(1)
arma21buffer(3)=@tdist(estmodel.@tstat(1),!dof)
arma21buffer(4)=@coefs(2)
arma21buffer(5)=@tstats(2)
arma21buffer(6)=@tdist(estmodel.@tstat(2),!dof)
arma21buffer(7)=@coefs(3)
arma21buffer(8)=@tstats(3)
arma21buffer(9)=@tdist(estmodel.@tstat(3),!dof)
arma21buffer(10)=@coefs(4)
arma21buffer(11)=@tstat(4)
arma21buffer(12)=@tdist(estmodel.@tstat(4),!dof)
arma21buffer(13)=@r2
arma21buffer(14)=@rbar2
rowplace(arma21parameters,@transpose(arma21buffer),!i)

!rvaverage=0
smirv_dummy=smirv
for !j=0 to !horizon-1
    smpl !i-!rolling_window+1+!j !i+!j
    equation estmodel.ls smirv_dummy c ar(1) ar(2) ma(1)
    forecastbuff-
er(!j+1)=estmodel.@coefs(2)*smirv_dummy(!i+!j)+estmodel.@coefs(3)*smirv_dummy(!i-
1+!j)+estmodel.@coefs(4)*resid(!i+!j)
    if forecastbuffer(!j+1)<0 then 'fix negative forecasts
endif
    smirv_dummy(!i+!j+1)=forecastbuffer(!j+1)
    !rvaverage=!rvaverage+forecastbuffer(!j+1)^2

```

```

next !j

smpl !i-!rolling_window+1 !i
rowplace(arma21forecast,@transpose(forecastbuffer),!i)
arma_30d(!i)=@SQRT((1/!horizon)*!rvaverage)

'CALCULATIONS FOR ARIMA(1,1,1)
'=====
equation estmodel.ls d(smirv) c ar(1) ma(1)
!dof= estmodel.@regobs-estmodel.@ncoef

arima111buffer(1)=@coefs(1)
arima111buffer(2)=@tstat(1)
arima111buffer(3)=@tdist(estmodel.@tstat(1),!dof)
arima111buffer(4)=@coefs(2)
arima111buffer(5)=@tstat(2)
arima111buffer(6)=@tdist(estmodel.@tstat(2),!dof)
arima111buffer(7)=@coefs(3)
arima111buffer(8)=@tstat(3)
arima111buffer(9)=@tdist(estmodel.@tstat(3),!dof)
arima111buffer(10)=@r2
arima111buffer(11)=@rbar2
rowplace(arima111parameters,@transpose(arima111buffer),!i)

!rvaverage=0
smirv_dummy=smirv

for !j=0 to !horizon-1
    smpl !i-!rolling_window+1+!j !i+!j
    equation estmodel.ls d(smirv_dummy) c ar(1) ma(1)
    forecastbuff-
er(!j+1)=smirv_dummy(!i+!j)+estmodel.@coefs(2)*(smirv_dummy(!i+!j)-smirv_dummy(!i-
1+!j))+estmodel.@coefs(3)*resid(!i+!j)
    if forecastbuffer(!j+1)<0 then 'fix negative forecasts by setting them
zero
        forecastbuffer(!j+1)=0
    endif
    smirv_dummy(!i+!j+1)=forecastbuffer(!j+1)
    !rvaverage=!rvaverage+forecastbuffer(!j+1)^2
next !j

smpl !i-!rolling_window+1 !i
rowplace(arima111forecast,@transpose(forecastbuffer),!i)
arima_30d(!i)=@SQRT((1/!horizon)*!rvaverage)

'CALCULATIONS FOR EGARCH
'=====
equation estmodel.arch(egarch) smilog c
estmodel.makegarch garchcondvar
!dof= estmodel.@regobs-estmodel.@ncoef

egarchbuffer(1)=@coefs(2)
egarchbuffer(2)=@tstat(2)
egarchbuffer(3)=@tdist(estmodel.@tstat(2),!dof)
egarchbuffer(4)=@coefs(3)
egarchbuffer(5)=@tstat(3)
egarchbuffer(6)=@tdist(estmodel.@tstat(3),!dof)
egarchbuffer(7)=@coefs(4)
egarchbuffer(8)=@tstat(4)
egarchbuffer(9)=@tdist(estmodel.@tstat(4),!dof)
egarchbuffer(10)=@coefs(5)
egarchbuffer(11)=@tstat(5)
egarchbuffer(12)=@tdist(estmodel.@tstat(5),!dof)
egarchbuffer(13)=@r2
egarchbuffer(14)=@rbar2
rowplace(egarchparameters,@transpose(egarchbuffer),!i)

!rvaverage=0

for !j=1 to !horizon
    if !j=1 then
        forecastbuff-
er(1)=@SQRT(@EXP(estmodel.@coefs(2)+estmodel.@coefs(3)*@ABS(resid(!i)/@SQRT(garchcondvar
(!i)))+estmodel.@coefs(4)*(resid(!i)/@SQRT(garchcondvar(!i)))+estmodel.@coefs(5)*@LOG(ga
rchcondvar(!i))))
    endif

```

```

        if !j>1 then
            forecastbuff-
er(!j)=@SQRT(@EXP(estmodel.@coefs(2)+estmodel.@coefs(3)*@ABS(forecastbuffer(!j-
1)/@SQRT(garchcondvar(!i))+estmodel.@coefs(4)*(forecastbuffer(!j-
1)/@SQRT(garchcondvar(!i))+estmodel.@coefs(5)*@LOG(forecastbuffer(!j-1)^2)))
            endif

            !rvaverage=!rvaverage+forecastbuffer(!j)^2
        next !j

        rowplace(egarchforecast,@transpose(forecastbuffer),!i)
        egarch_30d(!i)=@sqrt((1/!horizon)*!rvaverage)

next !i 'move the rolling_window by 1

'PREDICTIVE POWER
'=====
matrix(28,7) predictive_power=0
smpl !rolling_window !observations-!horizon

for !i=1 to 4
    if !i=1 then
        equation regr.ls(n) smirv_30d c arma_30d
    endif
    if !i=2 then
        equation regr.ls(n) smirv_30d c arima_30d
    endif
    if !i=3 then
        equation regr.ls(n) smirv_30d c egarch_30d
    endif
    if !i=4 then
        equation regr.ls(n) smirv_30d c vsmi_30d
    endif

    !dof= regr.@regobs-1
    predictive_power(!i*4-3,1)=@coefs(1)
    predictive_power(!i*4-2,1)=@stderrs(1)
    predictive_power(!i*4-1,1)=@tstat(1)
    predictive_power(!i*4,1)=@tdist(regr.@tstat(1),!dof)
    predictive_power(!i*4-3,!i+1)=@coefs(2)
    predictive_power(!i*4-2,!i+1)=@stderrs(2)
    predictive_power(!i*4-1,!i+1)=@tstat(2)
    predictive_power(!i*4,!i+1)=@tdist(regr.@tstat(2),!dof)
    predictive_power(!i*4-3,6)=@r2
    predictive_power(!i*4-3,7)=@rbar2
next !i

for !i=5 to 7
    if !i=5 then
        equation regr.ls(n) smirv_30d c arma_30d vsmi_30d
    endif
    if !i=6 then
        equation regr.ls(n) smirv_30d c arima_30d vsmi_30d
    endif
    if !i=7 then
        equation regr.ls(n) smirv_30d c egarch_30d vsmi_30d
    endif

    !dof= regr.@regobs-1
    predictive_power(!i*4-3,1)=@coefs(1)
    predictive_power(!i*4-2,1)=@stderrs(1)
    predictive_power(!i*4-1,1)=@tstat(1)
    predictive_power(!i*4,1)=@tdist(regr.@tstat(1),!dof)
    predictive_power(!i*4-3,!i-3)=@coefs(2)
    predictive_power(!i*4-2,!i-3)=@stderrs(2)
    predictive_power(!i*4-1,!i-3)=@tstat(2)
    predictive_power(!i*4,!i-3)=@tdist(regr.@tstat(2),!dof)
    predictive_power(!i*4-3,5)=@coefs(3)
    predictive_power(!i*4-2,5)=@stderrs(3)
    predictive_power(!i*4-1,5)=@tstat(3)
    predictive_power(!i*4,5)=@tdist(regr.@tstat(3),!dof)
    predictive_power(!i*4-3,6)=@r2
    predictive_power(!i*4-3,7)=@rbar2
next !i

```

```

'INFORMATION CONTENT
'=====
matrix(16,7) information_content=0

for !i=1 to 4
  if !i=1 then
    equation regr.ls(n) smirv_future c arma_30d
  endif
  if !i=2 then
    equation regr.ls(n) smirv_future c arima_30d
  endif
  if !i=3 then
    equation regr.ls(n) smirv_future c egarch_30d
  endif
  if !i=4 then
    equation regr.ls(n) smirv_future c vsmi_30d
  endif

  !dof= regr.@regobs-1

  information_content(!i*4-3,1)=@coefs(1)
  information_content(!i*4-2,1)=@stderrs(1)
  information_content(!i*4-1,1)=@tstat(1)
  information_content(!i*4,1)=@tdist(regr.@tstat(1),!dof)
  information_content(!i*4-3,!i+1)=@coefs(2)
  information_content(!i*4-2,!i+1)=@stderrs(2)
  information_content(!i*4-1,!i+1)=@tstat(2)
  information_content(!i*4,!i+1)=@tdist(regr.@tstat(2),!dof)
  information_content(!i*4-3,6)=@r2
  information_content(!i*4-3,7)=@rbar2
next !i

'BECKER
'=====
smpl @all

for %y rv_1d rv_5d rv_10d rv_15d rv_20d rv_25d drv_1d drv_5d drv_10d drv_15d
drv_20d drv_25d drv_30d 'defining variables for rv and drv
  series(!observations-!horizon) smi{%y}
next

matrix(!observations-!horizon,7) rv=0
matrix(!observations-!horizon,7) drv=0

for !i=!rolling_window to !observations-!horizon 'calculate RV and dRV

  !buffer=0
  for !j=1 to !horizon
    !buffer=!buffer+smirv(!i+!j)^2
    if !j=1 then
      smirv_1d(!i)=@SQRT(!buffer)
      smidrv_1d(!i)=@SQRT(!buffer)-smirv(!i)
    endif
    if !j=5 then
      smirv_5d(!i)=@SQRT(!buffer/5)
      smidrv_5d(!i)=@SQRT(!buffer/5)-smirv(!i)
    endif
    if !j=10 then
      smirv_10d(!i)=@SQRT(!buffer/10)
      smidrv_10d(!i)=@SQRT(!buffer/10)-smirv(!i)
    endif
    if !j=15 then
      smirv_15d(!i)=@SQRT(!buffer/15)
      smidrv_15d(!i)=@SQRT(!buffer/15)-smirv(!i)
    endif
    if !j=20 then
      smirv_20d(!i)=@SQRT(!buffer/20)
      smidrv_20d(!i)=@SQRT(!buffer/20)-smirv(!i)
    endif
    if !j=25 then
      smirv_25d(!i)=@SQRT(!buffer/25)
      smidrv_25d(!i)=@SQRT(!buffer/25)-smirv(!i)
    endif
    if !j=30 then
      'smirv_30d has already been calculated above

```

```

                                smidrv_30d(!i)=@SQRT(!buffer/30)-smirv(!i)
                                endif
                                next !j
                                next !i

matrix(36,5) becker_rv=0
matrix(36,5) becker_drv=0
smp1 252 2802

'RV
for !i=1 to 9
    if !i=1 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv
    endif
    if !i=2 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_30d smirv_25d smirv_20d smirv_15d
smirv_10d smirv_5d smirv_1d
    endif
    if !i=3 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_30d
    endif
    if !i=4 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_25d
    endif
    if !i=5 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_20d
    endif
    if !i=6 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_15d
    endif
    if !i=7 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_10d
    endif
    if !i=8 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_5d
    endif
    if !i=9 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smirv_1d
    endif

    !dof1= exinf.@regobs-1

    for !j=1 to 5
        becker_rv(!i*4-3,!j)=exinf.@coefs(!j)
        becker_rv(!i*4-2,!j)=exinf.@tstat(!j)
        becker_rv(!i*4-1,!j)=@tdist(exinf.@tstat(!j), !dof1)
    next !j
    becker_rv(!i*4,2)=exinf.@regobs*exinf.@jstat
    if !i=1 then
        becker_rv(!i*4,2)=0 'does not exist
    endif
    if !i=2 then
        becker_rv(!i*4,4)=1-@cchisq(becker_rv(!i*4,2),12-5)
    endif
    if !i>2 then
        becker_rv(!i*4,4)=1-@cchisq(becker_rv(!i*4,2),1)
    endif

'dRV
    if !i=1 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv
    endif
    if !i=2 then
        equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_30d smidrv_25d smidrv_20d smidrv_15d
smidrv_10d smidrv_5d smidrv_1d
    endif

```

```

        if !i=3 then
            equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_30d
        endif
        if !i=4 then
            equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_25d
        endif
        if !i=5 then
            equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_20d
        endif
        if !i=6 then
            equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_15d
        endif
        if !i=7 then
            equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_10d
        endif
        if !i=8 then
            equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_5d
        endif
        if !i=9 then
            equation exinf.gmm(b=a, n) vsmi_30d c arma_30d arima_30d egarch_30d smirv
@ c arma_30d arima_30d egarch_30d smirv smidrv_1d
        endif

        !dof1= exinf.@regobs-1

        for !j=1 to 5
            becker_drv(!i*4-3,!j)=exinf.@coefs(!j)
            becker_drv(!i*4-2,!j)=exinf.@tstat(!j)
            becker_drv(!i*4-1,!j)=@tdist(exinf.@tstat(!j), !dof1)
        next !j
        becker_drv(!i*4,2)=exinf.@regobs*exinf.@jstat
        if !i=1 then
            becker_drv(!i*4,2)=0 'does not exist
        endif
        if !i=2 then
            becker_drv(!i*4,4)=1-@cchisq(becker_drv(!i*4,2),12-5)
        endif
        if !i>2 then
            becker_drv(!i*4,4)=1-@cchisq(becker_drv(!i*4,2),1)
        endif
    next !i
    smpl @all

```

References

- Akgiray, V. (1989). Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts. *Journal of Business*, 62, 55-80.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96, 42-55.
- Barndorff-Nielsen, O.E., Shephard, N. (2004). Econometric analysis of realized covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica*, 72, 885-925.
- Becker, R., Clements A. E. & White S. I. (2007). Does implied volatility provide any information beyond that captured in model-based volatility forecasts? *Journal of Banking & Finance*, 31, 2535-2549.
- Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics*, 3, 167-179.
- Black, F. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81, 637-655.
- Brailsford, T.J. & Faff, R.W. (1996). An evolution of volatility forecasting technique. *Journal of Banking and Finance*, 20, 419-438.
- Brooks, C. (2008). *Introductory Econometrics for Finance*. Cambridge: Cambridge University Press.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- Bouchaud, J-P., Matacz, A. & Potters, M. (2001). Leverage Effect in Financial Markets: The Retarded Volatility Model. *Physical Review Letters*, 87:22, 228701.
- Campbell, J.Y., Lo, A.W. and MacKinlay A.C. (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.
- Canina, L. & Figlewski, S. (1993). The Informational Content of Implied Volatility. *Review of Financial Studies*, 6, 659-682.
- Christensen, B.J. and Prabhala, N.R. (1998). The Relation between Implied and Realized Volatility. *Journal of Financial Economics*. 50:2, 125–50.
- Christoffersen, P.F. and Diebold, F. (1987). Optimal prediction under asymmetric loss. *Econometric Theory*, 13, 808-817.

- Chu, K.-Y. (1978). Short-run Forecasting of Commodity Prices: An Application of Autoregressive Moving Average Models. *IMF Staff Papers*, 25, 90-111.
- Connolly, R. & Stivers, C. (2005). Macroeconomic news, stock turnover, and volatility clustering in daily stock returns. *The Journal of Financial Research*, 28, 235-259.
- Diebold, F. X., Hickman, A., Inoue, A. & Schuermann, T. (1998). Scale models. *RISK Magazine*, 11, 104-107.
- Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation. *Econometrica*, 50, 987–1008.
- Enders, W. (2004). *Applied Econometric Time Series*. New York: Wiley.
- Figlewski, S. (2004). *Forecasting Volatility*. Working Paper, New York University Stern School of Business.
- Fleming, J., Ostdiek, B. & Whaley, R.E. (1995). Predicting Stock Market Volatility: A New Measure. *Journal of Futures Markets*, 15:3, 265–302.
- Fleming, J. (1998). The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance*, 5:4, 317–345.
- Franses, P.H. & Van Dijk, D. (1996). Forecasting Stock Market Volatility Using (Non-Linear) GARCH Models. *Journal of Forecast*, 15, 229–35.
- Glosten, L. Jagannathan, R. & Runkle, D. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48, 1779–1801.
- Harvey, D. & Newbold, P. (2000). Tests for multiple forecast encompassing. *Journal of Applied Econometrics*, 15, 471-482.
- Hull, J. C. (2009). *Options, Futures and other Derivatives*. New Jersey: Pearson Education.
- Hwang, S., Knight, J.L & Satchell, S.E. (1999) *Forecasting Volatility Using LINEX Loss Functions*. Working Paper, University of Western Ontario.
- Lamoureux, C.G. & Lastrapes W.D. (1993). Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities. *Review of Financial Studies*, 6, 293-326.
- Liu, W., Morley, B. (2009). Volatility Forecasting in the Hang Seng Index using the GARCH Approach. *Asia-Pacific Financial Markets*, 16, 51-63.
- Mandelbrot, B. B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394.

- Makridakis, S. (1993). Accuracy measures: Theoretical and practical concerns. *International Journal of Forecasting*, 9, 527-529.
- Markowitz, H. (1952, March). Portfolio Selection. *Journal of Finance*, 7/1, 77-91.
- Nelson, D.B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59, 349-370.
- Newey, W.K., & West, K.W. (1987). Hypothesis Testing With Efficient Method of Moments Estimation. *International Economic Review*, 28:3, 777-788.
- Perry, P. (1982). The Time-Variance Relationship of Security Returns: Implications for the Return-Generating Stochastic Process. *Journal of Finance*, 37, 857-870.
- Pong S., Shackleton M.B., Taylor S. & Xu X. (2004). Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models. *Journal of Banking & Finance*, 28, 2541-2563.
- Poon (2005) *Forecasting Volatility in Financial Markets: A Practical Guide*. New York: Wiley.
- Poon, S.-H. & Granger, C.W.J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41, 478-539.
- Rubinstein, M. (1985). Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976, through August 31, 1978. *Journal of Finance*, 40, 455-480.
- SIX Swiss Exchange Ltd. (2009). *Guide Governing Volatility Index VSMI®*. Downloaded on April 10th 2010 from http://www.six-swiss-exchange.com/index_info/online/share_indices/smi/smifamily_factsheet_en.pdf
- SIX Swiss Exchange Ltd. (2010a). *SMI® Family Factsheet*. Downloaded on April 10th 2010 from http://www.six-swiss-exchange.com/index_info/online/other_indices/vsmi/vsmi_factsheet_en.pdf
- SIX Swiss Exchange Ltd. (2010b). *VSMI® Factsheet*. Downloaded on April 10th 2010 from http://www.six-swiss-exchange.com/index_info/online/other_indices/vsmi/vsmi_factsheet_en.pdf
- Steiner, M., & Bruns, C. (2002). *Wertpapiermanagement: Professionelle Wertpapieranalyse und Portfoliostrukturierung*. Stuttgart: Schäffer-Poeschel.
- Taylor, J.W. (2001). *Volatility Forecasting with Smooth Transition Exponential Smoothing*. Working paper, Oxford University.
- Taylor, S. (1986). *Modelling Financial Time Series*. New York: Wiley.

- Theil, H. (1966). Applied economic forecasting. Amsterdam: North-Holland Publ. Co.
- Theodossiou, P. (2000), *Skewed Generalized Error Distribution of Financial Assets and Option Pricing*. Working Paper, Rutgers Business School.
- Varian, H.R. (1975). *A Bayesian Approach to Real Estate Assessment*, in Fienberg, S.E. and Zellner, A. (ed.), *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J.S.*, Amsterdam: North-Holland.
- Yu, J. (2002). Forecasting volatility in the New Zealand stock market. *Applied Financial Economics*, 12, 193-202.
- Zellner, A. (1986). Bayesian Estimation and Prediction Using Asymmetric Loss Function. *Journal of American Statistical Association*, 81, 446-451.

List of figures

Figure 1 - Volatility smile for equities	9
Figure 2 - Implied distribution and lognormal distribution for equity options	10
Figure 3 - Estimation periods	19
Figure 4 - SMI and VSMI.....	27
Figure 5 - SMI daily realized volatility – time series view	27
Figure 6 - SMI daily realized volatility - descriptive statistics.....	28
Figure 7 - SMI 30-day realized volatility average - time series view	28
Figure 8 - SMI 30-day realized volatility average - descriptive statistics	29
Figure 9 - Initial ARMA(2,1) parameters.....	31
Figure 10 - Initial ARIMA(1,1,1) parameters	32
Figure 11 - Initial EGARCH parameters	33
Appendix Figure 1 - RV, VSMI and MBFs - Graph view	iv
Appendix Figure 2 – RV, VSMI and MBFs - Distributions	v

List of tables

Table 1 - Realized volatility - Average descriptive statistics*	30
Table 2 - Realized Volatility - Augmented Dickey-Fuller tests*	30
Table 3 - Estimated ARMA(2,1) parameters*.....	31
Table 4 - Estimated ARIMA(1,1,1) parameters*	32
Table 5 - Estimated EGARCH parameters*	33
Table 6 – Average 30-day forecast results*	34
Table 7 - Forecasting Quality	35
Table 8 - Forecasting Quality – LINEX	36
Table 9 - Predictive power analysis.....	37
Table 10 - Information content analysis	38
Table 11 - GMM estimation – SMIRV	39
Table 12 - GMM estimation – SMIDRV	40

List of used abbreviations

ADF	Augmented Dickey-Fuller test
AMAPE	Adjusted Mean Absolute Percentage Error
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
EGARCH	Exponential General Autoregressive Conditional Heteroscedasticity
EWMA	Exponential Moving Average
GARCH	General Autoregressive Conditional Heteroscedasticity
GJR-GARCH	The Glosten-Jagannathan-Runkle GARCH
GMM	Generalized Method of Moments
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
OLS	Ordinary Least Squares
OSMI	Best bid and best ask of all SMI-Options
RMSE	Root Mean Squared Error
RV	Realized Volatility
SIX	SIX Group
SMI	Swiss Market Index
VDAX-NEW	Volatility Index for the German Stock Index DAX
VIX	Chicago Board Options Exchange Volatility Index
VSMI	Volatility Index for the Swiss Market Index
VSTOXX	A market estimate of volatility on the EURO STOXX 50 Index
WSJ	Wall Street Journal