Forecasting the AEX Volatility Index

by
Rudy Smeets 698850

A thesis submitted in a partial fulfillment of the requirements for the degree of Bachelor in econometrics and Operations Research

Faculty of Economics and Business Administration
Tilburg University

Supervised by:
dr.F.C. Drost

Date:
August 31, 2010
Abstract

Several models are compared in the context of forecasting the AEX volatility index. Two approaches are considered in this thesis. The first approach is a univariate time series analysis of the VAEX. The second approach uses ARCH models estimated from the AEX index returns to forecast the AEX volatility index. The univariate model provides the most accurate forecasts of the VAEX.
Contents

1 Introduction 3

2 Data 5
   2.1 AEX Volatility index 5
   2.2 AEX Index log returns 8

3 Modelling the VAEX 9
   3.1 Univariate time series models 10
   3.2 Induced Volatility models 10
   3.3 Comparison to VAEX 13

4 Estimation methods 14
   4.1 Parameter estimation 14
   4.2 Rolling Window 15

5 Empirical results 16
   5.1 Results ARIMA model 16
   5.2 Results Volatility models 18
   5.3 performance forecasting VAEX 20

6 Conclusion 22
1 Introduction

In finance volatility is a measure for uncertainty and is an important input for a wide range of applications, such as asset pricing, risk management and asset allocation. For example consider the asset allocation theory by Markowitz (1952) in which he introduces a direct tradeoff between expected returns and volatility.

In practice it is not possible to directly observe the volatility of a certain financial instrument. The AEX volatility index (VAEX) measures the market’s expectation of volatility of the AEX for the next 30 days. This index is obtained by analyzing AEX option prices. This thesis examines several models for forecasting the VAEX. A logical approach is presented by Verbeek (2008). It proposes the use of historic data from the VAEX to estimate an univariate time series model.

A different approach is use volatility models based on AEX return data to forecasts values of the VAEX. Classic economic theory assumed that equity returns are normally distributed with a mean return and constant variance. Mandelbrot (1963) and Fama (1965) showed that the variance of speculative prices such as those of equity change over time. In this thesis we will consider flexible models that allows for changes in variance.

In financial time series one often observes volatility clustering. There are periods in which the volatility is high, such as during a crisis, and there are periods in which the volatility is low. As a result of this, big shocks on the market are often followed by big shocks in either direction. A prominent way to model volatility using the property of volatility clustering is the Autoregressive Conditional Heteroskedasticity (ARCH) model introduced by Engle (1982). The ARCH model was later generalized by Bollerslev (1986).

A disadvantage of the (G)ARCH model is that they are symmetric. A positive shock, good news, has the same impact on the volatility as a negative shock, bad news. However in equity markets a negative shock increase volatility more than an equal positive shock. An asymmetric model was developed by Glosten, Jagannathan and Runkle (1993), the GJR-GARCH.

It is shown in the previous studies by Bollerslev (1986) and Glosten et al (1993) that the parameters of the different volatility models are highly significant in sample and provide a good estimation of volatility. There is less evidence that the ARCH models provide good forecasts. An alternative is
Blair et al. (2001) show that forecasting with implied volatilities obtained from the S&P 100 Volatility Index performed better in forecasting realized volatility of the S&P 100. The ARCH models will be used to forecast volatility for a 30 day period. The value obtained from this procedure will be the forecast for the VAEX. We will compare all the models based on their ability to forecast the VAEX. This is done by measuring the forecast errors, the performance measures used are mean absolute deviations and the mean squared errors.

The thesis is arranged as follows. Section 2 discusses the data of AEX index and AEX Volatility Index. Section 3 presents an univariate model and several volatility models named in the introduction and their extensions. Section 4 discusses methods for estimation. In section 5 empirical results are presented. In section 6 we will discuss our conclusions.
2 Data

There are two data sets used in this thesis. The first set is a time series taken from the AEX volatility index, the second set is a time series taken from the AEX index. Both series consist of daily observations from the period 3 January 2000 up to 16 April 2010. The total number of observations is 2628.

2.1 AEX Volatility index

The AEX Volatility index (VAEX) is designed to measure the market’s expectation of volatility for the next 30 days. The measurement of the market’s expectation of volatility is done by analyzing option prizes. The volatility obtained from this type of analysis is called implied volatility.

For example consider the well known option pricing model given by Black and Scholes (1973). They show that under strong assumptions the price of an option depends on the current stock price, \( S_t \), the strike price, \( K \), time to expiration, \( T \), risk free rate, \( r \), and a constant volatility \( \sigma \).

The Black-Scholes model for a call-option is given by:

\[
F(S,t) = S\phi(d_1) - e^{-r(T-t)}K\phi(d_2)
\]

Where \( \phi(.) \) is the cumulative standard normal distribution function, and \( d_1 \) and \( d_2 \) are functions of \( S \) and \( t \) given by:

\[
d_1 = \frac{\log(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}
\]

\[
d_2 = \frac{\log(S/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}
\]

In the market we can observe the stock price, risk free rate and option price, the strike price and time to expiration are given by the option contract. Inserting all these inputs allows us to derive the implied volatility \( \hat{\sigma} \). However in practice the strong assumptions used for validating the Black-Scholes model are usually not met. To illustrate this we will give two examples.

First note that option contracts are standardized, the contracts have fixed rounded strike prices, and expire on the third friday of the expiry month. Deman and Kani (1994) show that options with different strike prices and the same time to expiration, return different implied volatilities. This effect is known as the volatility smile/smirk. The Black-Scholes model assumes
that a stock has a unique volatility, the volatility smile/smirk should not be possible in that setting. The Black and Scholes model also assumes that volatility is independent of the time to expiration. Deman and Kani (1994) show that the implied volatility increases as the time to expiration of the option contract increases, this implies that there is a dependence between implied volatility and time to expiration.

The AEX Volatility index is constructed in the same way as the CBOE (Chicago Board Option Exchange) volatility index (VIX). It is a weighted average of near-the-money put and call options. This is done to mitigate the effects of the volatility smile/smirk. The bias induced by time to expiration is mitigated by only weighting contracts that are due for expiration and expire next term. If the contracts are due within one week, the VIX rolls one term ahead to second and third term. This is done to avoid any pricing anomalies that might occur close to expiration. The exact calculation of the volatility index is beyond the scope of this thesis, however it is explained in the VIX methodology introduced by the Chicago Board Option Exchange.

In Figure 1 a graph of the data is presented, see Table 1 for the descriptive statistics of the AEX Volatility Index. The VAEX is designed to measure the market’s expectation of 30-day volatility implied by AEX index option prices and is quoted on an annualized basis. Figure 1 shows that volatility

![Figure 1: AEX Volatility Index for period 03/01/2000 until 16/04/2010](image)

is not constant. There are several periods in which the volatility is high and periods in which the volatility is low. The phenomena is called volatility
<table>
<thead>
<tr>
<th></th>
<th>AEX Volatility Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Volatility</td>
<td>25.54%</td>
</tr>
<tr>
<td>Maximum Volatility</td>
<td>81.22%</td>
</tr>
<tr>
<td>Minimum Volatility</td>
<td>10.12%</td>
</tr>
</tbody>
</table>

Table 1: Descriptive statistics for AEX Volatility index for period 03/01/2000 until 16/04/2010

clustering, this is common in financial time series. Table 1 shows that there is big difference between maximum and minimum volatility. The maximum volatility of 81.22% may seem unlikely since stocks cannot drop more than 100%. However if the 81.22% is transformed to the daily volatility by dividing it with \( \sqrt{256} \), it returns a daily volatility of 5.08%, which is not unlikely during a financial crisis.

Note that \( VAEX = \hat{\sigma} \times 100 \), we will denote the variance implied by the VAEX by \( VAEX^2 = \left( \frac{VAEX}{100} \right)^2 \). The VAEX is quoted on an annualized basis, for convenience we scale VAEX\(^2\) to daily and monthly (30 day period) denoted by \( VAEX^2_d = \frac{VAEX^2}{256} \) and \( VAEX^2_m = \frac{VAEX^2}{12} \).
2.2 AEX Index log returns

The second data set used is a time series of prices from the AEX index. The prices are used to determine the daily index log return of the AEX. This is done in the standard way by taking the natural logarithm of the ratio of two consecutive days. Let $S_t$ denote the price of the AEX at time $t$, the daily index return is given by $r_{aex} = \log S_t - \log S_{t-1}$.

In Figure 2 a graph of the Price Index is presented, the daily index return is shown in Figure 3, see Table 2 for descriptive statistics of the log index returns. In Figure 3 we notice that there are several periods with big shocks, as between 2002-2003 and 2008-2009 and periods with relative small shocks like in 2004-2007. This again indicates there is volatility clustering. Another observation we can make is that the periods with big shocks coincide with an overall decline in the AEX price index, this indicates that negative shocks might have more effect on the volatility than positive shocks.

Table 2 shows that the average return in the last decade was about -6.36% per year. The yearly volatility is 26.10%, which is higher that the market’s expected volatility, which is 25.54%, it seems that the market underestimates the volatility or the procedure used to construct the VAEX might be biased. However since the difference is relatively small we will ignore it. The maximum observed daily return was 10.03% and the minimum was -9.59%. The kurtosis is quite large compared to a standard normal distribution; this indicates that the returns do not follow a normal distribution.
Figure 3: Daily log returns of the AEX index for period of 03/01/2000 until 16/04/2010

<table>
<thead>
<tr>
<th>Daily log return AEX index</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average yearly log return</td>
<td>-6.38%</td>
</tr>
<tr>
<td>Standard Deviation yearly log return</td>
<td>26.10%</td>
</tr>
<tr>
<td>Maximum daily log return</td>
<td>10.03%</td>
</tr>
<tr>
<td>Minimum daily log return</td>
<td>-9.59%</td>
</tr>
<tr>
<td>Standard Deviation daily log return</td>
<td>1.61%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.07</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>9.02</td>
</tr>
</tbody>
</table>

Table 2: Descriptive statistics for daily log returns of the AEX index for period of 03/01/2000 until 16/04/2010

3 Modelling the VAEX

The goal of this thesis is to predict the value of the VAEX at time t+1 using only the information available at time t. In this section two different methods are explained that allow us to forecast the VAEX. The first method uses an univariate time series approach. The future value of the VAEX depends only on its past values. The time series mode is explained in section 3.1.

For the second method use the fact that the VAEX is the market’s expectation of volatility of the AEX index over the next 30 days. We use volatility models to analyze the volatility of the AEX index. With this information we can forecast the volatility of the AEX over a 30 day period. Aggregating the
forecasts of a 30 day period leads to an estimation of the VAEX. In section 3.2 we explain several different volatility models used to model the volatility of AEX. In section 3.3 we explain how to create multi period forecasts from the volatility models and how this leads to an estimation of the VAEX.

3.1 Univariate time series models

In this section an analysis of the VAEX is made using a univariate time approach. The future values of the VAEX depends (directly or indirectly) only on its past values. In reality the VAEX may depend on other economic quantities, however when it comes down to predictions we will not know the predictions of the other quantities rendering the model for the VAEX inadequate. Therefore we will use an univariate time series model.

In Verbeek (2008) the autoregressive integrated moving average (ARIMA) class of models is proposed. First we test if the series VAEX\(^2\) is stationary, by using a Dickey-Fuller test. If the series is stationary than it’s not integrated, otherwise the series has to be transformed by taking the difference between successive days, such that the difference series is stationary. Furthermore we will analyze the (partial) autocorrelation function to select the appropriate amount of ARMA terms and finally the Akaike information criterion is optimized to obtain a parsimonious efficient model. The empirical results are shown in section 5.

3.2 Induced Volatility models

In this subsection several volatility models are explained. All models use the AEX index returns to estimate the volatility. In section 3.3 a procedure is derived to link forecasts based on these volatility models with the VAEX.

Classic economic theory assumed that returns were normal distributed with constant mean and constant variance. A simple way to model volatility under the assumption of constant variance is by looking at the historic volatility. Let \( r_{aex} \) denote the returns of the AEX, the information available given by \( F \) consists of all the observations from time \( t = 0 \ldots T \), then the historic volatility \( \sigma_{hv} \) is given by:

\[
\sigma_{hv}^2 = \frac{1}{T-1} \sum_{i=1}^{T} (r_{aex,i} - \bar{r}_{aex})^2 \quad \text{with} \quad \bar{r}_{aex} = \frac{1}{T} \sum_{i=1}^{T} r_{aex,i}
\]

Although this gives an unbiased estimation of the sample variance, it is not sure that the future variance will be equal. Mandelbrot (1963) and Fama
(1965) show that variances change over time. We assume that the variance of the AEX changes over time, therefore we will only consider models that allow for changing variances.

**Generalized Autoregressive Conditional Heteroskedasticity** A popular way to model changing and clustered variances is Autoregressive Conditional Heteroskedasticity (ARCH) model of Engle (1982). It was later generalized by Bollerslev (1986) to the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. In the GARCH model the variance depends on its history. We assume that the return of the AEX index is constant over time, and denote this by $\mu$.

\[ r_{aex,t} = \mu + \epsilon_t \]

\[ \epsilon_t = v_t h_t^{\frac{1}{2}} \]

\[ v_t \text{ i.i.d, } E[v_t] = 0, \text{ var}(v_t) = 1 \]

Let $\epsilon_t$ denote the shocks on top of the average return $\mu$ and $\epsilon_t$ is a discrete time stochastic process. $h_t$ is a positive function, measurable with respect to the information set $\mathcal{F}$, available up to time t-1 and $h_t$ changes over time. $h_t$ denotes the variance $\sigma^2$ at time t. The GARCH(1,1) process is specified by

\[ h_t = E[\epsilon_t^2 | \mathcal{F}_{t-1}] = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} \]  

(1)

To ensure that $h_t \geq 0$ irrespective of $\epsilon_t^2$, we need to impose that $\omega, \alpha$ and $\beta$ are nonnegative. The formulation maintains the effect of volatility clustering, large shocks at t-1 increase the variance, making a large shock at time t more likely. The process $\epsilon_t$ is stationary for $\alpha + \beta < 1$. In financial time series we often find values close to 1, this implies a high persistence in volatility. Under stationarity $E[\epsilon_{t-1}^2] = E[h_{t-1}^2] = \sigma^2$. Verbeek (2008) shows that the unconditional variance of $\epsilon_t$ can be written as:

\[ \sigma^2 = \frac{\omega}{1 - \alpha - \beta} \]

The GARCH(1,1) can be rewritten to

\[ h_t = \sigma^2 + \alpha (\epsilon_{t-1}^2 - \sigma^2) + \beta (h_{t-1} - \sigma^2) \]

We can recursively substitute lags of the above in itself, resulting into:

\[ h_t - \sigma^2 = \alpha \sum_{j=1}^{\infty} \beta^{j-1} (\epsilon_{t-j}^2 - \sigma^2) \]
This shows that the variance in GARCH(1,1) depend on all the shocks up to time t. With $\beta < 1$ it implies that the effect of a shock on volatility diminishes over time, in practice shocks older than 2-3 years don’t have a significant effect on the variance.

An important restriction to the ARCH and GARCH models is their symmetry. The sign of the shocks do not matter, only their absolute values. A big negative shock has the same impact on future volatility as an equal sized positive shock. However in practice a negative shock may induce more uncertainty than a positive shock. An extension to this is allowing for asymmetry. Glosten et al (1993) developed the GJR-GARCH model, given by

$$h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-1}^2 I_{t-1} + \beta h_{t-1}$$

With $I_{t-1} = 1$ if $\epsilon_{t-1} < 0$ otherwise 0. If $\alpha_2 > 0$ then negative shocks have a larger impact on future volatility than positive shocks. Stationarity holds if $\alpha_1 + E[I] \alpha_2 + \beta < 1$, note that $E[I] = \frac{1}{2}$ if the returns have a symmetric distribution. Under stationarity the unconditional variance is given by:

$$\sigma^2 = \frac{\omega}{1 - \alpha_1 - E[I] \alpha_2 - \beta}$$

The unconditional variance of the GARCH model should be equal to the one of GJR-GARCH model.

**Combined Volatility models** Garch models work well for estimating variance in sample, however there is less evidence regarding good forecasting out of sample. Blair, Poon and Taylor (2001) proposed using implied volatilities from the VIX for forecasting the volatility of the S&P 100. We will add the implied volatility from the VAEX as exogenous variable to the GARCH models to test whether this improves the forecasts. For modelling purposes we transform the values of VAEX to $\text{VAEX}^2 = (\frac{\text{VAEX}}{100})^2/256$. The extended GARCH model becomes:

$$h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta h_{t-1} + \gamma \text{VAEX}^2_{t-1}$$

The GJR-GARCH is extended to:

$$h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-1}^2 I_{t-1} + \beta h_{t-1} + \gamma \text{VAEX}^2_{t-1}$$

under stationarity the unconditional variance if given by:

$$\sigma^2 = \frac{\omega + \gamma E[\text{VAEX}^2]}{1 - \alpha_1 - E[I] \alpha_2 - \beta}$$
3.3 Comparison to VAEX

The VAEX measures the market’s expectation of volatility for the next 30 days. In the previous section several models were introduced capable of forecasting the volatility of the AEX using the data of the returns. These models are used to create forecasts for the next 30 days, aggregation leads to a value that will serve as a proxy for VAEX\(_{t+1}\). For the GARCH and GJR-GARCH models, the one day ahead forecast is defined by the formulas:

\[
h_{t+1} = \omega + \alpha \epsilon_t^2 + \beta h_t \tag{5}
\]

\[
h_{t+1} = \omega + \alpha_1 \epsilon_t^2 + \alpha_2 \epsilon_t^2 I_t + \beta h_t \tag{6}
\]

The multi period ahead forecast is defined by the recursive formula:

\[E[h_{t+j}|F_t] = \omega + \rho E[h_{t+j-1}|F_t], \quad j > 1\]

with \(\rho\) is the persistence level, in the case of GARCH \(\rho = \alpha + \beta\) and in case of the GJR-GARCH \(\rho = \alpha_1 + E[I] \alpha_2 + \beta\). Aggregation of \(h_{t+j}\) leads to a proxy for VAEX\(_{t+1}\):

\[
\hat{VAEX}_{t+1}^2 = \sum_{j=1}^{21} h_{t+j} \tag{7}
\]

The formulation changes for the combined volatility models. There is no information about how the VAEX develops after \(t\). Therefore we assume that \(VAEX_{t+j} = VAEX_t\). The one period ahead forecast is defined by:

\[
h_{t+1} = \omega + \alpha \epsilon_t^2 + \beta h_t + \gamma VAEX_t^2 \tag{8}
\]

\[
h_{t+1} = \omega + \alpha_1 \epsilon_t^2 + \alpha_2 \epsilon_t^2 I_t + \beta h_t + \gamma VAEX_t^2 \tag{9}
\]

The multi period ahead forecast is defined by the recursive formula:

\[E[h_{t+j}|F_t] = \omega + \rho E[h_{t+j-1}|F_t] + \gamma VAEX_t^2, \quad j > 1\]

again \(\rho\) is the persistence level, in the case of extended GARCH \(\rho = \alpha + \beta\) and in case of the extended GJR-GARCH \(\rho = \alpha_1 + E[I] \alpha_2 + \beta\). Equation 7 is used to calculate the proxy for VAEX\(_{t+1}\).
4 Estimation methods

In the first section we will explain how to obtain parameter estimates for the ARMA model and (GJR)-GARCH models with/without the exogenous variable. In the second section the rolling window procedure is explained, which is used to create a series of proxies for the VAEX.

4.1 Parameter estimation

Estimation is done by optimizing the log likelihood function. This requires an assumption about the distribution of $\epsilon_t$, commonly normality is assumed. If $\epsilon_t$ does not have a standard normal distribution, then the maximum likelihood procedure may still provide consistent estimators for the model parameters. Under weak assumptions the first-order conditions of the maximum likelihood procedure are also valid when $\epsilon_t$ is not normally distributed. This is referred to as quasi-maximum-likelihood.

First we will derive the maximum likelihood function for the volatility models. Note that $\epsilon_t = y_t - \mu$ and $\epsilon_t = \nu_t h_t^{\frac{1}{2}}$. In the general case $h_t$ is a strict positive function of $z_t' = (1, \epsilon_{t-1}^2, \epsilon_{t-1}^2 I_{t-1}, \hat{\epsilon}_{t-1}, \text{VAEX}_{t-1})$ and $\theta' = (\omega, \alpha_1, \alpha_2, \beta, \gamma)$ such that $h_t = z_t' \theta$. We can estimate the GARCH(1,1) model by restricting $\alpha_2$ and $\delta$ to zero, the other models can be estimated by making similar restrictions. To obtain a likelihood function we assume that the conditional density of $\nu_t$ is the standard normal distribution, so that $\epsilon_t | \mathcal{F}_{t-1} \sim N(0, h_t)$. Bollerslev (1986) shows that the log likelihood function for a sample of $T$ observations is:

$$L_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} l_t(\theta)$$

$$l_t = -\frac{1}{2} \log h_t - \frac{1}{2} \frac{\epsilon_t^2}{h_t^{\frac{1}{2}}}$$

Optimizing the log likelihood function with respect to $\theta$ leads to the parameter estimations. The likelihood function for the ARMA model is similar. $\epsilon_t$ is more complicated and is defined by the ARMA structure. $h_t$ equals the unconditional variance $\sigma^2$, since in ARMA model the $\epsilon_t | \mathcal{F}_{t-1} \sim N(0, \sigma^2)$. Optimizing over the ARMA parameters leads to the estimates.
4.2 Rolling Window

The goal is to compare the models mentioned in section 3 on their ability to forecast the VAEX. A rolling window of information $F_t$ is used to create a series of proxies of the VAEX, these will be compared on their average error. The total dataset consists out of the 2628 observations. The window length is set to 3 years which corresponds to the first 768 observations. This produces the first forecast of the VAEX, the one on 2-jan-2004. Next the window rolls one day to produce a forecast for 3-jan-2004 using the observations 2 up to 769, etc. Old observation are deleted to make the estimations more flexible. The procedure is repeated until the end of the sample is reached, resulting in a series of 1860 forecasts of the VAEX per model.
5 Empirical results

This section presents the empirical results obtained from analyzing the data. In the first subsection we present and discuss the results of ARIMA model. The second subsection presents and discusses the results of the volatility models. In the last sub section a comparison is made of the forecasting performance of all models.

5.1 Results ARIMA model

We apply an univariate time series analysis as proposed in section 3. The VAEX series is transformed to the implied variance by \( \text{VAEX}^2 = \left( \frac{\text{VAEX}}{100} \right)^2 \). First we test the series for stationarity by applying the Dickey-Fuller test. This leads to a Dickey-Fuller test with intercept and 6 lags, with a t-Statistic of -4.00. This rejects the presence of a unit root at the 1% level. Based on this test we conclude that the series is stationary.

Next we use a graphical inspection of the (partial) autocorrelation function to select the appropriate number of AR and MA terms. The (P)ACF is shown in figure 4. The ACF remains significant for a long time, however the PACF is not significant after a some lags. This suggest a model with only AR terms. The Schwarz information criterion (SIC) is optimized to obtain a parsimonious efficient model. The SIC was optimal for an AR(6) model. For simplicity we denote \( \text{VAEX}^2_t \) by \( Y_t \) and \( y_t = Y_t - E[Y_t] = Y_t - \mu \). The model is specified as followed:

\[
y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \theta_3 y_{t-3} + \theta_4 y_{t-4} + \theta_5 y_{t-5} + \theta_6 y_{t-6} + \epsilon_t
\]

Results of the AR(6) model using the first 3 years as data are shown in Table 3. Only \( \mu \) and \( \theta_1 \) are significant on a 5% level. The other variables are
Estimates AR(6) model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\mu$</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\theta_3$</th>
<th>$\theta_4$</th>
<th>$\theta_5$</th>
<th>$\theta_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
<td>0.0084</td>
<td>1.0903</td>
<td>-0.1724</td>
<td>0.0623</td>
<td>0.0859</td>
<td>-0.1278</td>
<td>0.0457</td>
</tr>
<tr>
<td>t-ratio</td>
<td>(2.72)</td>
<td>(12.33)</td>
<td>(-1.77)</td>
<td>(0.48)</td>
<td>(0.76)</td>
<td>(-1.27)</td>
<td>(0.72)</td>
</tr>
</tbody>
</table>

Table 3: Results AR(6) based on first 3 years of data

not. However these are the results of just one period. Since we use a rolling window there are many (different) estimates. The results of this is shown in Table 4. The rolling window method adds flexibility to the estimates, at the cost of interpretation. Only the mean and the first lagged variable have a clear positive impact, the other 5 lagged variables don’t have clear positive or negative impact on the forecast of the VAEX.

<table>
<thead>
<tr>
<th>Stability of AR(6) model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
</tbody>
</table>

Table 4: Descriptive Statistics of all the estimates generated by the rolling window for the AR(6) model
5.2 Results Volatility models

Table 5 presents the parameter estimates, robust t-ratios log likelihoods and unconditional variances, for the 4 volatility models defined in Section 3. The results are based on the first 3 years of the data, from January 2000 to December 2003.

The first model is the standard GARCH(1,1) model. All the parameters are significant on a 5% level, the persistence is close but smaller than one with $\alpha_1 + \beta = 0.9905$. The GARCH(1,1) model is the most parsimonious of all four models and has the lowest log likelihood score. However the three remaining models have a higher or equal log likelihood score by definition since they extend the GARCH(1,1) model. The second model is the asymmetric GJR-GARCH model. $\alpha_2$ is significant, this implies that negative shocks have a different impact on variance than positive shocks. The persistence is with $\alpha_1 + E[I] \alpha_2 + \beta = 0.9938$ similar to the first model. The third model is the GARCH(1,1) extended with an exogenous variable. Both $\beta$ and $\gamma$ are not significant on a 10% level. This suggests there is an identification problem between $h_{t-1}$ and $VAEX_{t-1}^2$. The fourth model is the extended GJR GARCH model. $\alpha_2$ is significant, just like in model 2 this implies that negative shocks have a different impact on the variance than positive shocks. $\beta$ and $\gamma$ are now both significant, adding an asymmetry component solved the problem of identification.

Under stationarity all models should have the same unconditional variance. In Table 5 we see that the unconditional variance are close to each other. The sample variance of this period is $3.3050 \times 10^{-4}$, the small deviations from this can be explained by estimation errors.
Daily log index returns modelled by the arch specification:

\[ r_{aex,t} = \mu + \epsilon_t \]
\[ \epsilon_t = h_t^{\frac{1}{2}} z_t, \quad z_t \sim i.i.d. N(0,1) \]

(1) \[ h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta h_{t-1} \]
(2) \[ h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-1}^2 I_{t-1} + \beta h_{t-1} \]
(3) \[ h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta h_{t-1} + \gamma V_{AEX} t_{t-1} \]
(4) \[ h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-1}^2 I_{t-1} + \beta h_{t-1} + \gamma V_{AEX} t_{t-1} \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model (1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega \times 10^{-6} )</td>
<td>3.4622</td>
<td>1.9885</td>
<td>-60.59</td>
<td>-9.2163</td>
</tr>
<tr>
<td></td>
<td>(2.01)</td>
<td>(3.03)</td>
<td>(1.56)</td>
<td>(7.73)</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.1223</td>
<td>-0.0427</td>
<td>-0.0584</td>
<td>-0.0975</td>
</tr>
<tr>
<td></td>
<td>(3.09)</td>
<td>(-2.84)</td>
<td>(-2.73)</td>
<td>(-5.43)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.2243</td>
<td>0.2590</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.07)</td>
<td>(7.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.8682</td>
<td>0.9255</td>
<td>0.2361</td>
<td>0.8444</td>
</tr>
<tr>
<td></td>
<td>(23.50)</td>
<td>(67.85)</td>
<td>(0.45)</td>
<td>(30.44)</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>0.8915</td>
<td>0.1374</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.56)</td>
<td>(7.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-L</td>
<td>2142.45</td>
<td>2168.67</td>
<td>2170.88</td>
<td>2183.26</td>
</tr>
<tr>
<td>Excess Log-L</td>
<td>26.22</td>
<td>28.43</td>
<td>40.81</td>
<td></td>
</tr>
<tr>
<td>UV* ( \sigma^2 \times 10^{-4} )</td>
<td>3.6274</td>
<td>3.2457</td>
<td>3.2845</td>
<td>3.3437</td>
</tr>
</tbody>
</table>

*Unconditional Variance Notes: the parameters are estimated by maximizing the quasi-log-likelihood functions, assuming normal densities. Robust t-ratios are shown in parentheses.

Table 5: Models for the AEX index, daily returns from January 2000 to December 2003

Table 6 present descriptive statistics of the estimations as they occurred after finishing the rolling window procedure. The GARCH(1,1) model is the most stable model, the parameters show little variations. In contrary to the extended GARCH+VAEX all parameters change a lot for different periods.
\[ h_t = \omega + \alpha_1 \epsilon^2_{t-1} + \alpha_2 \epsilon^2_{t-1} h_{t-1} + \beta h_{t-1} + \gamma \text{VAEX}^2_{t-1} \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
<th>Model (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega \times 10^{-6} )</td>
<td>2.6376</td>
<td>2.0367</td>
<td>-31.854</td>
<td>-0.2525</td>
</tr>
<tr>
<td>Mean</td>
<td>1.0818</td>
<td>0.7517</td>
<td>28.698</td>
<td>3.0161</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.1023</td>
<td>-0.0318</td>
<td>0.0253</td>
<td>-0.0556</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0278</td>
<td>0.0179</td>
<td>0.0622</td>
<td>0.0195</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.1644</td>
<td>0.2098</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.1644</td>
<td>0.2098</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.1644</td>
<td>0.2098</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.0413</td>
<td>0.0492</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.8792</td>
<td>0.9332</td>
<td>0.0734</td>
<td>0.8757</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.0313</td>
<td>0.0162</td>
<td>0.3973</td>
<td>0.0424</td>
</tr>
<tr>
<td>Mean</td>
<td>0.8068</td>
<td>0.0507</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.4303</td>
<td>0.0340</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Descriptive Statistics of all the estimates generated by the rolling window for the volatility models

5.3 performance forecasting VAEX

The goal of the thesis is to forecast the VAEX. For each model we obtain a time serie of forecasts, using the method of rolling window as explained in section 4. As proxy we use \( \text{VAEX}^2 \) quoted on a monthly basis, \( \text{VAEX}^2 = \frac{\text{VAEX}}{100}^2 / 12. \) In the case of the univariate time series model \( \hat{\text{VAEX}}^2 = \sum_{i=1}^{21} h_{t+i} \), we aggregate only 21 days, since a month has 21 trading days. The forecasting error, \( e \), is defined by the difference between the actual value of \( \text{VAEX}^2_m \) and the forecast \( \hat{\text{VAEX}}^2_m \).

The models will be compared on basis of their forecast errors. The models with the smallest errors will be the best performing models. There are two methods used for measuring the errors. The first is Mean Absolute Deviations:

\[ \text{MAD} = \frac{1}{n} \sum_{i=1}^{n} |e| \]
the second method is Mean Squared Errors:

\[ MSE = \frac{1}{n} \sum_{i=1}^{n} e^2 \]

large errors get relatively more weight using the MSE instead of MAD. The results of these tests are shown in table 7.

Table 7 shows that the AR(6) model has the best performance on forecasting the VAEX\(^2\) and the GJR-GARCH model performs slightly better than the GARCH model, it has both a lower MAD and a lower MSE. But when these models are extended by the VAEX, the extended GARCH model performs better. The extended models perform much better than the normal models, this indicates that the AEX index doesn’t contain all the information about the VAEX.
6 Conclusion

This thesis examined several models for forecasting the VAEX. The standard univariate time series model has the smallest prediction errors. All the proposed volatility models made larger prediction errors. This may be caused by the difficulties of forecasting the volatility of the AEX itself. Since the VAEX is constructed by a weighted average of AEX options, there might be a more complex link between the real volatility and the VAEX than proposed in this thesis.
References


E Derman, I Kani, 1994 Riding on a smile, Risk 1994


