Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Trading Volume and Volatility in the US Stock Markets

Author: **Tomáš Juchelka**

Supervisor: **PhDr. Boril Šopov, Msc., LL.M.**

Year of defence: **2014**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 16, 2014

Signature

Acknowledgment

First of all, I would like to thank my academic supervisor PhDr. Boril Šopov, Msc., LL.M from the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, for the guidance, suggestions, comments and knowledge he provided that helped me writing this thesis.
I am also deeply indebted to Jan Juchelka, Petr Skotnica and Josef Němý for helping me with the data issues.
Last but not least, I owe thanks to all those who supported me during the work on the thesis.

Bibliography Reference

Juchelka, Tomas. *Trading Volume and the Volatility in the US Stock Markets* Prague, 2014. 42 p. Bachelor thesis (Bc.) Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor: PhDr. Boril Šopov, Msc., LL.M.

Extent of the Thesis

81 696 (with spaces)

Abstract

This thesis investigates the relationship between trading volume and stock return volatility using GARCH model in the framework of Mixture of Distribution Hypothesis. Analysis is carried out for five well-known stocks selected from the American S&P500 stock index. Our approach was to extend the variance equation of the well known GARCH model on the trading volume which was split into three explanatory variables capturing different effects of volume on volatility. Apart from the relationship itself, we examined the changes of GARCH and ARCH parameters after the inclusion of volume, implicitly testing the Mixture of Distribution Hypothesis. Interesting results and implications for future research were identified. Firstly, we highlight the appropriateness of the volume decomposition into expected and unexpected volume, where all the volume parameters turned out to be statistically significant. General observation was that the increase of both expected and unexpected trading volume leads to the increase of volatility. On the other hand, negative volume shocks tend to decrease it. Eventhough we performed the analysis with lagged and also contemporaneous volume, we were not able to confirm that the inclusion of volume leads to insignificance of the ARCH and GARCH parameters, thus not confirming the Mixture of Distribution Hypothesis. However, we found that the volume models perform significantly better than the plain GARCH models according to AIC. Considering these findings, it is possible to conclude that there is positive relationship between the stock return volatility and trading volume. We also found that the volume models perform substantially better in modeling and predicting the future volatility.

JEL Classification C22, C52, C55, G12
Keywords GARCH, volatility, trading volume, Mixture of Distribution Hypothesis

Author's e-mail tom.juchelka@gmail.com
Supervisor's e-mail boril.sopov@gmail.com

Abstrakt

Tato práce zkoumá vztah mezi počtem zobchodovaných akcií a volatilitou při použití GARCH modelu v Hypotéze Mixu Distribucí. Analýza je provedena na pěti známých akciích, vybraných z indexu S&P500. Našim přístupem bylo obohacení druhé rovnice GARCH modelu o dodatečné vysvětlující proměnné, které reprezentovaly počet zobchodovaných akcií. Kromě samotného vztahu volatility a objemu obchodů jsme zkoumali, jaký vliv má zahrnutí dodatečných proměnných na statistickou významnost GARCH a ARCH parametrů v našem modelu. Zjistili jsme, že rozdělení objemu obchodů na očekávanou a neočekávanou složku bylo namístě, jelikož se všechny tyto proměnné ukázaly být statisticky významné. Našim hlavním zjištěním bylo, že vzrůst počtu zobchodovaných akcií, tedy vzrůst očekávaného a neočekávaného objemu, vede ke zvýšení volatility. Na druhou stranu negativní objemový šok, tedy snížení neočekávaného objemu, vede k jejímu snížení. Přestože jsme naši analýzu provedli jak pro první lag objemu obchodu, tak také pro současný objem, nepodařilo se nám potvrdit hypotézu, že zahrnutí objemu obchodů způsobí statistickou bezvýznamnost parametrů ARCH a GARCH, tedy nepotvrdili jsme Hypotézu Mixu Distribucí. Zjistili jsme však, že modely se zahrnutým objemem jsou podstatně lepší, než běžné modely typu GARCH. Výsledky naší analýzy by se daly shrnout závěrem, že s rostoucím počtem obchodů roste volatilita a tedy existuje statisticky významný pozitivní vztah volatility a počtu zobchodovaných akcií. Dále je třeba říci, že modely rozšířené o objem jsou podstatně lepší, než běžné modely typu GARCH.

Klasifikace JEL C22, C52, C55, G12

Klíčová slova GARCH, volatilita, objem obchodů, Hypotéza Mixu Distribucí

Email autora tom.juchelka@gmail.com

E-mail vedoucího práce boril.sopov@gmail.com

Contents

List of Tables

List of Figures

Acronyms

MDH Mixture of Distributions Hypothesis

SIAH Sequential Information Arrival Hypothesis

ARCH Autoregressive Conditional Heteroscedastic

GARCH Generalized Autoregressive Conditional Heteroscedastic

AR Autoregressive

ARMA Autoregressive Moving Average

EGARCH Exponencial GARCH

AIC Akaike Information Criterion

BIC Bayesian Information Criterion

TIC Takeuchi's Information Criterion

K-L Kullback-Leibler Distance

ACF Auto Correlation Function

Bachelor Thesis Proposal

Author	Tomáš Juchelka
Supervisor	PhDr. Boril Šopov, Msc., LL.M.
Proposed topic	Trading Volume and Volatility in the US Stock Markets

Topic characteristics The work is predominantly going to be concerned about the trading volume and volatility in the US stock markets, basically figuring out the relationship between this two key concepts in finance. I choose this topic because there are extensive discussions about the volatility and I find it important to shed light on the relationship between volatility and the trading volume, as one of the causes. I will try to construct rigorous model which could help us to investigace the relationship and comparing the results obtained from the empirical analysis with the theory of finance and our expectations. In the first part, the work is going to be concerned about the volatility modeling theory, followed by theoretical concepts that are to explain the relationship. In the second part we will conduct the empirical research, figuring out the estimation results and confronting the results with our expectations.

Outline

1. Introduction
2. Theoretical Framework
    2.1. Conditional Heteroscedastic Models
    2.2. Model Selection Method
    2.3. Volume Volatility Relationship
3. Empirical Research
    3.1. Data and Methodology
    3.2. Model Application
    3.3. Summary

4. Conclusion

Core bibliography

1. Bollerslev, T (1986): "Generalized autoregresive conditional heteroscedastic ." *Journal of Econometrics* **31**: pp. 307-327.
2. Clark, P.K (1973): "A Subordinated Stochastic Process Model with Finite Variance For Speculative Prices." *Econometrica* **41(1)**: pp. 135–155.
3. Nelson, D. (1991): "Conditional heteroscedasticity in asset returns: A new approach." *Econometrica* **59(2)**: pp. 347–370.
4. Lamoureux, Christopher G. & W.D. Lastrapes (1990): "Heteroscedasticity in Stock Return Data: Volume versus GARCH Effects." *The Journal of Finance* **45(1)**: pp. 221–229.
5. Tsay, Ruey S., Wiley(2005): *"Analysis of Financial Time Series"* 2nd Edition, John Wiley & Sons, Inc. ISBN: 978-0-471-74618-8

Author

Supervisor

Chapter 1

Introduction

There has been extensive research in the field of quantitative finance and financial econometrics regarding the key financial markets variables such as stock returns, trading volume, volatility or bid-ask spreads resulting in many interesting findings. Despite the fact that the non-normality of returns was empirically observed much earlier, it was firstly scientifically described by famous mathematician Benoit Mandelbrot in the early 1960's in his work called *The variation of certain speculative prices* Mandelbrot (1963). Namely, he found that returns exhibit excess kurtosis which is that either big positive or big negative returns occure more often than normal distribution would imply. He also observed clustering of return volatility, where periods of high volatility are followed by periods of high volatility and periods of low volatility are followed by periods of low volatility. The volatility clustering played important role in econometric research that led to the formation of GARCH family models. Developed by Engle (1982) and later generalized by Bollerslev (1986), the Autoregressive Conditional Heteroscedastic Engle (1982) and Generalized Autoregressive Conditional Heteroscedastic Bollerslev (1986) models proved to be exceptional among other models in modeling and predicting the return volatility, since they capture both the clustering and non-normality of returns. While other models at that time operated under the assumption of constant variance, the ARCH and GARCH processes allow the conditional variance to change over time, leaving the unconditional variance constant. In fact, constant volatility is rather uncommon in empirical evidence. Financial markets also exhibit that volatility is higher in falling market than in rising market. This asymmetrical behaviour was firstly documented by Black (1976) and led to the number of improvements of GARCH models, namely, EGARCH Nelson (1991), Threshold GARCH model Zakoian (1994) and GJR-GARCH model by L.R. Glosten (1993). Since volatility is key imput variable in many financial models, the precise modeling is of great importance in econometrics and should be done carefully.

Encouraged by fairly valid modeling of volatility, researchers started to put the volatility in context of prior information flows which included variables such as trading volume, returns or bid-ask spreads. Particularly, there is a lot of research in the area of volume-volatility relationship with various interesting interpretations. There are suggestions in the markets that high trading volume is associated with above average volatility, since the prices need volume to change. Moreover, the volume tend to be higher with positive returns rather than negative returns. These suggestions were confirmed by many empirical researches for example by Ali F. Darrat (2003) who examined intraday data of sellected stocks in the Dow Jones Industrial Average and found that high trading volume casues high volatility. The scientific process also led to the formation of theories explaining the volatility foundations including volume-volatility relationship. We can find two competing theories explaining the volume-volatility relationship, the theory of information flows and dispersion of beliefs theory. The theory of iformation flows consists of two hypothesis, the Mixture of Distributions Hypothesis (MDH) developed by Clark (1973) and Sequential Information Arrival Hypothesis (SIAH) developed by Copeland (1976).

The Mixture of Distributions Hypothesis says that both volume and volatility are variables which are jointly determined by serially corelated mixing variable. Such mixing variable measures the rate at which information arrive to the markets. Moreover the MDH assumes that all market participants receive the information simultaneously and the price shifts directly to the new equilidbrium level, thus both the volume and volatility change contemporaneously vis-a-vis the mixing variable. According to the MDH, researchers should also be able to explain the volatility persistance using the mixing variable. For example Lamoureux (1990) have examined the volume-volatility relationship and found supportive evidence in favour of MDH.

Another hypothesis that links the volume-volatility relationship with information flow is the previously mentioned sequential information arrival hypothesis. According to Copeland (1976), traders receive information in random order changing their trade position accordingly to new information, thereby, creating the volume and volatility. However, the information flows to the market sequentially, therefore, not all traders have the information available simultaneously. Accordingly the price is gradually shifting towards equilibrium which is established when the information is received by the whole market. As a result, there should be evidence on both lag and contemporaneous relationships between volume and volatility.

The second theory concerned about the volume-volatility is so called dispersion of beliefs theory. It states that each trader attach different importance to each information and trades on it, thus it leads to above average volume and volatility. The theory also incorporate the distinction between informed, and uninformed trading, its reactions to new informations and its effects on volume and volatility.

In our work, we plan to conduct the research on volume-volatility relationship in terms of MDH, thus apart from examining the relationship itself, we will focus on the magnitude and significance of the GARCH and ARCH parameters. We decided to choose top five US stocks using five minute return and volume data. The stocks will be selected according to the market capitalisation, where we plan to include five companies with the highest market capitalisation in the S&P500 stock market index.

Our intention is to augment one of the GARCH family models on the trading volume, moreover, the trading volume will be split into expected and unexpected part, capturing different fundaments of trading volume and showing their effects on the stock return volatility. On the top of that, we want to take a look at the volatility persistance parameters which, according to MDH, should become statistically insignificant after the inclusion of trading volume. In the end, we will compare the "explanatory" ability of the plain GARCH models with those with volume included in the variance equation, where we expect that the inclusion of trading volume will improve the explanatory power of the model.

The thesis is organized as follows: Chapter 2 establishes the theoretical framework, where the GARCH models are gradually introduced and foundations of our research are set. In Chapter 3 the data and methodology are described, later followed by the model application section, broken down into five categories according to the particular stock. In Chapter 4 the final conclusion is provided along with the propositions for future research.

Chapter 2

Theoretical Framework

2.1 Conditional Heteroscedastic Models

The objective of this section is to expand the theory around the topic of volatility and its modeling which shall be later put in context of Mixture of Distributions Hypothesis, we also plan to introduce the model which we use in our empirical research.

Since the asset volatility is an important input in many econometric models, let us begin with its definition followed by some important charasteristics that have been seen in financial markets and are crucial for our analysis.

Under the term of asset volatility, we understand the conditional standart deviation of the underlying asset returt which measures the rate of return fluctuation in the markets, thus serve also as the risk indicator. Asset volatility has many financial applications, besides the options pricing models, portfolio theory and risk management, the proper modeling of volatility can also improve the efficiency of parameter estimation and accuracy in interval forecasts of many financial time series models.

One of the main features of volatility is the fact that the volatility is not directly observable, since we observe only the realised returns. Newertheless, it has some characteristics that are commonly seen in asset returns noted in Ruey S. Tsay (2005) . Firstly, the volatility is not constant over time, but evolves in continuous manner which means we can rarely see volatility jumps. Secondly, the volatility clusters, where there are periods of low volatility and periods of high volatility. Thirdly, long memory property, where shocks from remote periods affect the current volatility. Fourthly, the volatility is stationary process, where it varies in fixed range which also means that longterm volatility predictions should converge to its long term average. Fifth, the asymmetric behaviour of volatility, where volatility reacts differently to big price drop and big price increase. This property is also known under the name of leverage effect.

Also the volume-volatility relationship plays important role in financial time series research. The main motivation for such modeling is the suggestion that volume can be used as a proxy for information flows to the market, however, the prior research results are contradicting. The inclusion of trading volume should also provide some new information to the volatility models, thus make the estimates more precise.

All these characteristics played important role in the evolution of volatility modeling and also contributed to the formation of Conditional Heteroscedastic models. In the following sections, we are going to provide you an insight in the ARCH process, introducing ARCH models, its generalisations and extensions that overcome the initial shortcomings.

At the end, the model will be put in context of the volume-volatility framework in terms of Mixture of Distributions Hypothesis. Last, the model we selected for volume-volatility modeling will be introduced.

2.1.1 The ARCH model

The first conditional heteroscedastic model that provides systematic approach for volatility modeling is the Autoregressive Conditional Heteroscedastic (ARCH) model developed by Engle (1982) later followed by number of improvements and generalisations.

Before introducing the ARCH process itself, it is useful to show the modeling of the serially uncorrelated but dependet return series r_t and its properties, namely the conditional mean and variance.
The return series is modeled as follows:

$$r_t = \mu + a_t \tag{2.1}$$

where conditional mean of r_t is $E(r_t|F_{t-1}) = \mu$ and conditional variance of r_t is given by $Var(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}] = h_t$. It is useful to notice that term F_{t-1} stands for the information set available at time $t-1$. The last term a_t that appears in our model is reffered to as the shock, or the inovation of return series at time t. It can be written as follows:

$$a_t = \sqrt{h_t}\epsilon_t$$

where $\{\epsilon_t\}$ is an i.i.d. sequence of random variables with zero mean and variance one.

The objective of an ARCH is to model the conditional variance h_t. Let us begin with the simple ARCH(1), later extendet to ARCH(q) model.
The ARCH(1) model is specified as:

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 \tag{2.2}$$

where again $a_t = \sqrt{h_t}\epsilon_t$, $\alpha_1 \geq 0$ and $\alpha_0 > 0$ are parameters to be estimated. The ARCH(1) model is just a special case of general ARCH(q) model which is specified as follows.

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \ldots + \alpha_q a_{t-q}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2 \tag{2.3}$$

Where again non-negativity conditions $\alpha_0 > 0$ and $\alpha_i \geq 0$ for all $i > 0$ must hold, in order to ensure non-negativity of conditional variance. Coefficients α_i must also satisfy some regularity conditions to obtain finite unconditional variance of a_t.

From the structure of the model we can see that large past squared shocks a_{t-i}^2 cause large conditional variance h_t of a_t which in ARCH framework means that large shocks tend to be followed by another large shock. This is similar to the volatility clustering feature of the asset returns, thus the ARCH(q) model is useful in modeling the financial time series.

Despite all its useful properties, ARCH model has also some shortcomings, namely the fact, that it responds similarily to positive and negative shocks, because conditional variance is dependent on the squares of a_t.

ARCH also does not provide any explanation of the source of variations, it just mechanically describes the conditional variance. Another disadvantage that the model is struggling with is the fact that it usually has to be specified with high order lag structure, since the long memory feature is often present in financial series data. There is also probability of violation of non-negativity conditions, when it comes to the estimation of the free lag distribution.

To overcome these problems, the model was generalised to GARCH which will be provided in the next section.

2.1.2 The GARCH model

Althought the ARCH model proved to be exceptional in volatility modeling, some improvements had to be done mainly due to issues mentioned above and also to allow for more flexible lag structure. In this section, more general class of processes, GARCH, is introduced.

As you may know, the GARCH model was suggested by Bollerslev (1986). The extension of the ARCH process to the GARCH bears much resemblance to the extension of the time series Autoregressive (AR) process to the general Autoregressive Moving Average (ARMA) process and allows us more carefull description in many situations.
The GARCH(p,q) process is given by following set of equations:

$$r_t = \mu + \epsilon_t = \mu + \sqrt{h_t} z_t \tag{2.4}$$

$$\epsilon_t | \Omega_{t-1} \sim N(0, h_t) \tag{2.5}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{2.6}$$

from equation 2.4 we can see that $\epsilon_t = r_t - \mu$ is real value, discrete time, strictly stationary stochastic process, z_t is independent identically ditributed standart normal variable, Ω_{t-1} is the information set of all information available through time $t-1$. Parameters p,q determine the number of lags included in the model, α_0, α_i, β_j are parameters to be estimated which are subjects of the following conditions: $\alpha_0 > 0$, $\alpha_i \geq 0$ for $i = 1 \dots q$ measuring the short-term impact of ϵ_t on conditional variance and $\beta_j \geq 0$ for $j = 1 \dots p$ measuring the long-term impact on conditional variance.

From the equation 2.6 is obvious that if $p = 0$ the GARCH reduces to the ARCH(q) process and if $p = q = 0$ ϵ_t is simply white noise. All in all, the most widely used specification of the GARCH asserts that the best predictor of the future variance is a weighted average of the long-run average variance, the variance predicted for current period, and the information in this period captured by the most recent squared residual. This kind of process corresponds to so called adaptive learning behaviour.

Along with the non-negativity conditions, the $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$ must be satisfied to ensure the covariance stationarity of GARCH(p,q) process. To put it down rigorously, let us provide you the following theorem provided by Bollerslev (1986) on page 310.

Theorem 1. The GARCH(p,q) process defined by 2.5 and 2.6 is wide-sense stationary with $E(\epsilon_t) = 0$ and $Var(\epsilon_t) = \alpha_0(1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j)^{-1}$ and $cov(\epsilon_t, \epsilon_s) = 0$ for all $t \neq s$ is and only if $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$

The GARCH framework proved to be successful in volatility prediction. Empirically, a wide range of financial phenomena exhibit the volatility clustering. As we have seen, the GARCH model describes the time evolution of average size of squared errors that is, the evolution of the magnitude of uncertainty.

But again, to stay objective, we must not forget to provide the downsides of the GARCH model. Nelson (1991) has mentioned three main disadvantages of the model. The non-negativity condition of the estimation parameters which are imposed in order to ensure the non-negativity of h_t in all time periods. This coditions lead to the fact that increasing ϵ_t in any time period leads to increase of h_{t+m} for all m $\geqslant$ t. This eliminates the randomness in the fluctuations of h_t. GARCH is also unable to capture the leverage effect, thus the magnitude of h_t does not deppend whether $\epsilon_t = r_t - u \leq 0$ or $epsilon_t = r_t - u \geq 0$, though it is observed that $\epsilon_t \leq 0$ is related to higher h_t and $\epsilon_t \geq 0$ is related to lover h_t. The last drawback that Nelson mentioned is the difficulty in the evaluation of persistance of shocks to conditional volatility.

To overcome these problems, the Exponencial GARCH or EGARCH was developed which we present in the subsequent section.

2.1.3 The Exponencial GARCH model

As we have already mentioned, Nelson (1991) has developed the Exponencial GARCH (EGARCH) model in order to overcome the issues that GARCH was unable to deal with. Mainly to allow for the asymmetric effect in asset returns. The EGARCH model can be specified by adding additional parameter to the modified GARCH equation which allows us for the asymmetric behaviour.
The EGARCH(p,q) model is defined as:

$$r_t = \mu + \epsilon_t = \mu_t + \sqrt{h_t} z_t \tag{2.7}$$

$$\epsilon_t|\Omega_{t-1} \sim N(0, h_t) \tag{2.8}$$

$$ln(h_t) = \omega + \sum_{i=1}^{p} \beta_i ln(h_{t-i}) + \sum_{j=1}^{q} \alpha_j g(z_{t-j}) \tag{2.9}$$

Where $z_t = \dfrac{\epsilon_t}{\sqrt{h_t}}$ is standartized residual, Ω_{t-1} is information set available through time $t-1$, $g(z_t)$ is a function of z_t and $|z_t|$ that accommodate the asymmetric effect between stock return and volatility changes.
The function is defined as follows:

$$g(z_t) = \theta z_t + \gamma[|z_t| - E|z_t|] \tag{2.10}$$

Both z_t and $[|z_t| - E|z_t|]$ are i.i.d., zero-mean sequences with continuous distributions. Since $g(z_t)$ is just the linear combination of the two, it is also i.i.d., zero-mean sequence, where θ and γ are real constants.

The asymmetric behaviour of $g(z_t)$ can be seen from the following equation:

$$g(z_t) = \begin{cases} (\theta+\gamma)z_t - \gamma E(|z_t|), & \text{if } z_t \geq 0 \\ (\theta-\gamma)z_t - \gamma E(|z_t|), & \text{if } z_t < 0 \end{cases}$$

Over the range of $0 < z_t < \infty$, $g(z_t)$ is linear function of z_t with the slope $(\theta + \gamma)$. Over the range $-\infty < z_t < 0$ $g(z_t)$ is linear function of z_t with the slope equal to $(\theta - \gamma)$, thus it allows the conditional variance h_t to respond asymmetrically to positive and negative shocks.

As you can see from the equation 2.9 the variance $ln(h_t)$ is in the logarithmic form which ensures the non-negativity of h_t in case of negativity of estimation parameters ω, β_i and α_j.

In order to properly understand the meaning of the estimation parameters, let us rewrite the equation 2.9 in the following manner:

$$ln(h_t) = \omega + \sum_{i=1}^{p} \beta_i ln(h_{t-i}) + \sum_{j=1}^{q} \lambda_j[|z_t| - E|z_t|] + \sum_{j=1}^{q} \delta_j z_{t-j} \tag{2.11}$$

where ω, β_i, λ_j, δ_j are parameters to be estimated.

Parameter ω is just a constant, β_i is conditional volatility persistance parameter, λ_j represents the symmetric GARCH effect and δ_j measures the asymmetric effect. It is good to notice that λ_j and δ_j parameters are nothing else than just transformations of the preceding estimation parameters such that $\lambda_j = \alpha_j \gamma$ and $\delta_j = \alpha_j \theta$.

As for the asymmetric effect, we can distinguish among three cases that may occure, $\delta > 0$ indicates that positive realisation of innovation has positive effect, $\delta < 0$ indicates that negative realisation of innovation results in increase of volatility and $\delta = 0$ is the case, when the asymmetric behaviour is not taken into account. Most often, in empirical aplications of the model, the parameter δ is assumed to be negative.

In the next sections of the text, we would like to explain the theoretical basis for the selection of the proper number of lags in our variance equation. The tool we are going to use is also useful for overall model comparison which can also serve for comparison of the GARCH models with and without volume terms.

2.2 Model Selection Method

Although we have already presented different forms of the GARCH model, we have not said anything specific about the lag distribution and the process by which will be the appropriate number of parameters p,q employed in our model.

Generaly, when fitting the model, it is possible to increase the likelihood just by including additional parameters into the model but doing so may result in the model overspecification and loss of the predictive ability of the model.

There exist number of selection approaches that deal with the model specification issue helping us to find the model that best fits the particular analysis. Each of those processes is suitable in different situation and uses different statistical method in the model evaluation process. The most widely utilized are for example Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), Takeuchi's Information Criterion (TIC) or second order information criterion.

Although all of these criterions mentioned above are useful in certain situations, for our purposes the AIC will be utilized, since it is usually used in case of models with predictive rather than descriptive ability and carries other useful theoretical properties outlined in Kenneth P. Burnham (2002).

Before we introduce Akaike Information Criterion itself, we provide a brief insight to its main building block, the Kullback-Leibler distance noted as K-L distance.

2.2.1 Kullback-Leibler distance

In 1951, Sollomon Kullblack and Richard Leibler, S. Kullblack (1951) derived an information measure now reffered to as the Kullback-Leibler (K-L) information which can be conceptualized as a directed distance between two models, say f and g. The f and g are known simple probability distributions, where f is a notation for full reality and g denotes an approximating model in terms of probability distribution. It is good to notice that the measure of f to g is not the same as the measure of g to f, since it can be conceptualized as a directed, or oriented distance.

Kullback-Leibler information between f and g for continuous functions is defined as the multi-dimensional integral

$$I(f,g) = \int f(x) log\left(\frac{f(x)}{g(x|\theta)}\right) \mathrm{d}x. \tag{2.12}$$

where log denotes natural logarithm. $I(f,g)$ denotes the information lost when f is approximated by g, thus can be understood as the distance from g to f.

Full reality f is considered to be fixed, whereas g varies over the space of models indexed by θ. Similarly, the K-L distance can be understood as the measure of inefficiency assuming that the distribution is g when the true distribution is f. The K-L distance is always positive, unless the two distributions f and g are identical such that $I(f,g) = 0$, if $f(x) = g(x)$. Note that the estimation of Kullback-Leibler distance requires both the knowledge of the true distribution f as well as all the parameters in the models g_i, thus the distance itself can not be computed for real world problems.

However, the problem of not-knowing the exact functions f and g drops out, as we use the relative distance, where $I(f,g)$ can be written as

$$I(f,g) = \int f(x) log\left(f(x)\right) \mathrm{d}x. - \int f(x) log\left(g(x|\theta)\right) \mathrm{d}x. \tag{2.13}$$

Where each of the integrals is a statistical expectation with respect to f. So the K-L distance can be rewritten into the following form.

$$I(f,g) = E_f\left[log(f(x))\right] - E_f\left[log(g(x|\theta))\right] \tag{2.14}$$

The fitst expectation $E_f\left[log(f(x))\right]$ is a constant that depends only on the unknown true distribution f, since we assume that f is unambigous and does not change. Therefore, treating the expectation as a constant, say C, allows us to measure the relative directed distance such that

$$I(f,g) = C - E_f\left[log(g(x|\theta))\right] \tag{2.15}$$

or

$$I(f,g) - C = -E_f\left[log(g(x|\theta))\right] \tag{2.16}$$

the term $I(f,g) - C$ can be interpreted as the measure of the relative distance between f and g. Assume that we have two models g_1 and g_2 such that $I(f,g_1) < I(f,g_2)$ which means that the first model is better than the secod. Than the following applies $I(f,g_1) - C < I(f,g_2) - C$, hence
$-E_f\left[log(g_1(x|\theta))\right] < -E_f\left[log(g_2(x|\theta))\right]$,
,moreover,
$I(f,g_2) - I(f,g_1) = -E_f\left[log(g_2(x|\theta))\right] + -E_f\left[log(g_1(x|\theta))\right]$,
so even without knowing C we can identify which model is better and how much better it is, because C is the same for all candicate models, and as you can see, it is irrelevant for comparison.

In the next section we will show how the relative K-L distance was used for building the information criteria that we use for model selection issues.

2.2.2 Akaike Information Criterion

In 1973, Akaike (1973) proposed the use of the Kullblack-Leibler information as a fundamental basis for model selection issues. As we know, the actual K-L information cannot be computed, unless we know the functions f and all the parameters θ in each function $g_i(x|\theta)$, thus Akaike suggested that relative K-L distance can be estimated based on the log-likelihood function at its maximum point.

In fact, he noted that there exists unknown unique value of θ which minimizes the K-L distance and depends on the function f, the structure of the model g, the sample space and the parameter space.

Since the model parameters must be estimated, the task is to minimize the expected the K-L distance rather than the known K-L distance. So Akaike proposed to estimate the $E_yE_x\left[log(g(x|\hat{\theta}(y)))\right]$ by the maximised $log(L(\hat{\theta})|data)$. He also found that this estimation is upward biased and under certain conditions, this bias is equal to K, the number of estimation parameters in the approximating model.

Based on this, the following equality was suggested:

$$\hat{E}(K-L) = E_y E_x \left[log(g(x|\hat{\theta}(y)))\right] = log(L(\hat{\theta})|data) - K \tag{2.17}$$

where $g()$ is the candidate model, L is the maximized likelihood function, K is number of parameters included in the approximating model, $(K-L)$ is the Kullback-Leibler distance and θ are the model parameters.

Later than, Akaike created the Akaike Information Criterion given by the following equation:

$$AIC = -2log(L(\hat{\theta})|data) + 2K \tag{2.18}$$

As we know, the $I(f,g)$ can be smaller when including additional known parameters in the approximating model g, but when the parameters are unknown and must be estimated further uncertainty is added to the estimation of the relative K-L distance, thus at some point, inclusion of additional parameters will have the opposite effect of increasing the estimated relative K-L distance.

This can be seen from the right hand side of the equation 2.18 where $-2log(L(\hat{\theta})|data)$ tends to decrease, when additional parameters are included, while the term $2k$ tends to get larger as more parameters are included, thus somehow penalises more complex and overspecified models.

At the end it is good to notice that AIC provide us rather tool for the comparison of the models included in the model set, than some absolute measure of the fit, so if we have a set of poor models, AIC will give us the best one, even though it may still be very poor in absolute sence.

2.3 Volume Volatility Relationship

Presenting conditional heteroscedastic models that properly capture the features of volatility, we made rather mechanical description of the problem than providing you the reasons and explanations of such behaviour. In this section, we will focus to correctly introduce the theory which deals with the explanation of volatility and its features through connection with trading volume.

2.3.1 Mixture of Distributions Hypothesis

A Mixture of Distributions Hypothesis (MDH) developed by Clark (1973) is being widely used for the explanation of volatility features through the connection of volatility and trading volume. For the beginning, let us provide you technical description of the matter and continuously build the entire theory.

Let us assume the daily stock return equation:

$$r_t = \mu + \epsilon_t$$

where again, μ is mean of r_t conditional on past information, ϵ_t is a shock to return series such that $\epsilon_t|\Omega_{t-1} \sim N(0, h_t)$, where Ω_{t-1} is the set of information available through time $t-1$.
The Mixture of Distributions Hypothesis suggest the following:

$$\epsilon_t = \sum_{i=1}^{n_t} \delta_{it}$$

where δ_{it} is the ith equilibrium price increment in day t, n_t is the random variable that denotes the number of daily increments representing the stochastic rate of the information arrival to the market. The sequence of δ_{it} is assumed to be the sequence of independently identicaly distributed random variables with zero mean and variance σ^2. Moreover, if n_t is sufficiently large, than $\epsilon_t|n_t \sim N(0, \sigma^2 n_t)$. Since the number of intraday increments is random, daily returns follow mixture of normally distributed random variables with n_t as the mixing variable. Thus, the daily returns are generated by subordinated stochastic process, where ϵ_t is subordinate to δ_i and n_t is the directing process.

According to Lamoureux (1990), GARCH effects may be explained as a manifestation of time dependence in the rate of evolution of intraday equilibrium returns. For the validity of the argument, let we assume the serial correlation in the daily number of information arrivals which can be expressed by subsequent equation:

$$n_t = \phi_0 + \sum_{i=1}^{q} \phi_i n_{t-i} + u_t \tag{2.19}$$

where ϕ_0 is a constant, ϕ_i are slope parameters of the laged n_t and u_t is white noise. Let us define $\Theta_t = E(\epsilon_t^2|n_t)$. If the mixture of distributions model is valid than $\Theta_t = \sigma^2 n_t$. From equation 2.19 plugged for n_t into $\Theta_t = \sigma^2 n_t$ follows:

$$\Theta_t = \sigma^2 \phi_0 + \sum_{i=1}^{q} \phi_i \Theta_{t-1} + \sigma^2 u_t \tag{2.20}$$

This equation illustrates the fundaments of Mixture of Distribution Hypothesis. It captures the same kind of persistance in conditional variance of returns that can be estimated by GARCH model, where the autoregressive form of the mixing variable n_t is translated into the GARCH model structure. As you can see, the conditional variance is dependent on n_t the number of information that flows to the market, thus GARCH behaviour is formed by serially correlated information flow process.

The unpleasant fact that the n_t is generally unobservable calls for the need of some suitable proxy variable. According to Lamoureux (1990), daily trading volume can be used. The volume is likely to contain the information about disequilibtium dynamics of the stock markets. It is important to notice that in case of contemporaneous relation, trading volume was assumed to be weakly exogenous.
Therefore, the following model was proposed:

$$h_t = \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} + \gamma V_t \tag{2.21}$$

which is simple GARCH(1,1) specification augmented on V_t the daily trading volume term which serves as the a proxy for the mixing variable.

The mixture of distribution suggests that $\gamma > 0$, moreover, the $(\alpha+\beta)$ which measure the persistance of conditional variance should become negligible because inclusion of trading volume should make parameters α and β statistically insignificant if the trading volume is serially correlated. Additional insight into the problem may provide us decomposition of volume into expected and unexpected part assuming that the market consists of different kind of traders.

$$h_t = \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} + \gamma_1 EV_t + \gamma_2 UV_t \tag{2.22}$$

Coefficients γ_1 and γ_2 measure the effect of expected and unexpected volumes on volatility. Usually it is observed that increase of unexpected volume results in higher volatility and increase of expected volume in lower volatility, thus we can expect corresponding signs of our estimation parameters in the regression.

At this part, it is necessary to present you one drawback that arises from inclusion of contemporaneous trading volume. As you can find above, MDH states that both trading volume and return volatility are driven by the information flow process. In other words, volume and volatility are both determined by the same variable, thus volatility can not be regressed with contemporaneous volume as exogenous variable due to simultaneity problem resulting in inconsistent estimators. Moreover, model with contemporaneous volume would not be capable of forcasting volatility.

The solution of simultaneity proposed by Najand (1991) is to include lagged trading volume V_{t-1} instead of contemporaneous V_t in the GARCH equation, where lagged values of endogenous variables are classified as predetermined, thus the lagged value can be used. Besides, there has been found a positive significant relationship between lagged volume and volatity. As mentioned earlier, Lamoureux (1990) showed that inclusion of contemporaneous trading volume leads to reduction of volatility persistance parameters in the GARCH equation, thus explaining the volatility. Chen (2001) came to the finding that the persistance is not eliminated after inclusion of lagged or contemporaneous trading volume, thus obtaining contradicting results. Later in 1994, Lamoureux (1994) performed the volume-volatility analysis again, this time, they developed the model the without the assumption of weak exogenity of trading volume resulting in quite opposite results than in previous analysis.

All these contradicting resuls led to further development of the problem ,namely to division of the trading volume into expected and unexpected components with the assumption that expected volume representing the liquidity trading should lead to decrease of volatility and unexpected volume representing the flow of information to the market should lead to increase of the volatility.

Overall, there exist rather inconclusive evidence in the matter of volume-volatility relationship and explanation of persistance parameters through inclusion of trading volume in the variance equation which motivates us for construction of suitable model for volatility dynamics in the presence of information arrival proxies.

2.3.2 Augmented GARCH

Until now, we have introduced basic theoretical concepts and econometric models suitable for utilisation of volume-volatility modeling and Mixture of Distribution Hypothesis testing. In this section, we would like to put all the pieces together leading to introduction of the model we use in our empirical research.

Since we want to capture the volatility as realistically as possible but still keep the interpretation simple and understandable for the broad audience of readers, we decided to use GARCH model which performs very well among all the models mentioned above. Despite it does not capture all the empirically observable features of volatility, namely the asymmetric behaviour of innovations, it will ensure the simple interpretation of estimation results of the augmented GARCH model. As for the additional explanatory variables added to the GARCH equation, we suggest that the most proper way would be to include the lagged volume rather than contemporaneous to avoid the potential simultaneity problems. We also decompose the volume into expected and unexpected parts rather than just using simple lagged volume. In addition to show the asymmetric impact of the volume change on the volatility, we decided to decompose the unexpected volume into two variables capturing the positive and negative shocks. The proposed model has the following form:

$$h_t = \omega + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{j=1}^{q} \alpha_j a_{t-j}^2 + \gamma_1 EV_{t-1} + \gamma_2 UVP_{t-1} + \gamma_3 UVN_{t-1} \quad (2.23)$$

where EV_{t-1} is expected volume in time $t-1$, UVP_{t-1} is a variable capturing the positive volume shocks in time $t-1$ and the UVN_{t-1} is a variable capturing negative volume shocks in time $t-1$.

This distinction of expected and unexpected trading volume has also interesting interpretation, where expected volume can be understood as liquidity trading, thus it should not have anything to do with the information arrival. On the other hand, the unexpected volume can be interpreted as volume occured due to new, unexpected news arriwal to the market. The expected signs of our estimation parameters are quite intuitive, since the liquidity trading is a volume that occure periodically and liquidity is assumed to be stabilizing factor of the markets than we can expect the negative sign of the expected volume parameter. As mentioned above, the unexpected volume represents the information trading, however, it can be positive or negative, thus we should account for these two options in our regression. The positive unexpected volume is assumed to increase the stock volatility and negative unexpected volume is assume to decrease it. The lag structure of the $\sum_{i=1}^{p} \beta_i h_{t-i}$ and $\sum_{j=1}^{q} \alpha_j a_{t-j}^2$ will be specified according to Akaike information criterion introduced in section 2.2.2..

Chapter 3

Empirical Research

3.1 Data and Methodology

3.1.1 Data description

As mentioned in the introduction, we decided to make a research on five S&P 500 constituents with the highest current market capitalisation. The data was obtained from the Bloomberg [1] system and processed in the Stata software. Generaly S&P 500 index represents the common stock of the 500 largest US companies publicly listed on either NYSE or NASDAQ exchange. S&P letters stand for Standard&Poor's which is an American financial services company that founded the index in 1957. All in all the importance of S&P500 lies in the fact that the index is being used as a benchmark for American stock performance and bellwether for the US economy.

As mentioned above, we have picked five companies with the highest market capitalisation to date 8th of April 2014.
The selected companies are as follows:

- Apple Inccorporated
- Exxon Mobil Corporation
- Berkshire Hathaway Incorporated
- Microsoft Corporation
- Google Incorporated

[1] We have to thank for the data to anonymous provider.

These stocks represent currently the most valuated companies in US economy, thus we will briefly present their history, background and business activities they perform.

Apple Inc. is an American technology corporation founded in 1976, headquartered in California. The company develops, designs and sells consumer electronics and software. Later in 1980, Apple launched the initial public offering of its stock for 22 dollars per share. Currently Apple inc. stocks are traded in NASDAQ stock exchange around 520 dollars per share with market capitalisation of 463 486 million dollars.

Exxon Mobil Corp. is an American Oil&Gas company, formed in 1999 by the merger of two major Oil companies Exxon and Mobil. The company is headquartered in Texas. Although the company was formed in 1999, the stocks of Exxon and Mobil were traded far earlier in twenties of the 20th century. Nowadays the stocks are traded in NYSE, fluctuating around 95 dollars per share and the market capitalisation is 416 895 million dollars.

Berkshire Hathaway Inc. is an American conglomeral holding company founded in 1839 headquartered in Nebraska USA, originally, Berkshire operated as textile manufacturing company but later in 1962 was acquired by Warren Buffet and expanded into insurance and other businesses. Nowadays the stocks of Berkshire Hathaway are traded in NYSE. Currently the price per share is around 180 dollars with market capitalisation of 300 281 million dollars.

Microsoft Corporation is an American technology company founded in 1975, headquartered in Washington, USA. The company mainly develops software and consumer electronics. In 1986, Microsoft firstly issued its shares to public for 21 dollars per share. Nowadays the Microsoft stocks are traded in NASDAQ around 40 dollars per share, the market capitalisation of Microsoft is 325 463 million dollars.

Google Inc. is the last and the youngest company we want to introduce. Google was founded in 1994 in California, it is a technology company that specialises on the internet and software services and solutions. Google initially issued its shares in 2004 for 85 dollars per share, currently the shares are trading around 550 dollars and the market capitalisation is 359 457 million dollars.

For all of these companies except Google, we collected raw intraday five minute data for the close price and aggregated volume for the period beginning from 1.10.2013 till 31.3.2014. Each regular trading day begins at 15:30:00 CET and ends at 22:00:00 CET which makes 78 consecutive observations per regular trading day. Moreover after exclusion of holidays, weekends and non-trading hours we obtained 9678 observations for each of the stocks. In a Google case, we had to restrict ourselves just for the period of 1.10.2013 - 7.3.2014 due to many missing ovservations beginning in 10.3.2014 and consecutive days. All in all we collected 8427 observations, however, this restriction generally should not cause any problems in our analysis. For each stock, period and number of observations is summarized in the table 3.1.

Table 3.1: Data collection summary

Company	From	Until	Number of obs.
Apple	1.10.2013	31.3.2014	9678
Exxon Mobil	1.10.2013	31.3.2014	9678
Berkshire Hathaway	1.10.2013	31.3.2014	9678
Microsoft	1.10.2013	31.3.2014	9678
Google	1.10.2013	7.3.2014	8427

Source: author's computations.

For those who are not familiar with terms such as close price or aggregated trading volume, the explanation is provided in the subsequent rows, where we also explain how the data were collected.

The close price is generally the latest price for which the security was traded in given time period, thus it represents the latest up-to-date valuation of the security. Since we use five minute periods in our analysis, the five minute close price is the latest price of the security for every period. As you know, we divided the trading day into 78 subsequent five minute periods and picked the close prices for each of the period. From these close prices, the five minute returns were calculated. In order to stay consistent with the periodicity of prices, we also had to collect the aggregated volume for each of the five minute interval which means that the volumes represent the sum of all traded stocks during the given five minute interval.

The lenght of the time series was chosen according to data availability, where the longest period supported by bloomberg platform was the six months for the five minute observations of close price and volume. However, the number of observations is sufficiently large for all stocks, thus we do not have to be worried about any problems connected with the quality of the estimation results.

Now we move on to the next section, where the method of computation and statistical properties of returns are discussed.

3.1.2 Return series

Intraday stock returns series introduces the added dificulty in our research. Due to the fact that we work with the five minute intraday data, the series is rather discontinuous. Because of existence of nontrading hours, weekends and holidays, there are many five minute intervals without any observation either of price or volume. By definition, we calculate the return series as the log-returns which is shown in the following equation.

$$r_t = ln\left(\frac{P_t}{P_{t-1}}\right) = ln(P_t) - ln(P_{t-1})$$

P_t is the close price in period t, P_{t-1} is previous period close price and ln is simply natural logarithm.

At this point, we should notice the problem of discontinuity of the time series. Since the returns are computed from this periods and previous five minute periods close price, we should somehow treat the fact that we dont have any previous five minute close price observation for the first observation of the opening period (15:30-15:35) of the day. Generally, there are two ways how to deal with this problem.

One solution is to proxy the opening price of the day as the previous day's closing price which implicitly means that we assume that the time between previous day closing price and today first close price observation is simply five minutes. Althought this approach gives us the advantage of pseudocontinuity of time series, it also has its disadvantages mentioned later this section. The second approach would be to give up the return of the first five minute period of the day and simply begin with the (15:35-15:40) as the first one for each of the trading days. Now, let us provide you the summary statistics of the return series with the data based on both of the approaches. Starting with the first one, we obtained the following figures. The results from the second approach, where we dropped each day's first observation are summarized in table 3.3 .

If we compare the values of the two, we can see decrease in all of the summary categories. The biggest differences can be noticed in minimum and maximum returns which can suggest us that the biggest absolute returns occur

Table 3.2: 5 minute returns summary statistic #1

Company	Mean	Std. Dev.	Min	Max
Apple	0.0000122	0.0015398	-0.0795977	0.028973
Exxon Mobil	0.0000131	0.0010869	-0.0192283	0.0220017
Berkshire Hathaway	9,93e-6	0.0008921	-0.0144388	0.0118229
Microsoft	0.0000214	0.0015731	-0.0146484	0.0640195
Google	0.0000388	0.0016412	-0.0111212	0.0972701

Source: author's computations.

Table 3.3: 5 minute returns summary statistic #2

Company	Mean	Std. Dev.	Min	Max
Apple	-6.97e-06	0.0010948	-0.014034	0.0062963
Exxon Mobil	0.0000132	0.0008974	-0.0079965	0.0085749
Berkshire Hathaway	2.68e-07	0.0007895	-0.0144388	0.0052429
Microsoft	3.25e-06	0.0011888	-0.0146484	.0103723
Google	7.01e-06	0.00106	-0.0088511	0.0090204

Source: author's computations.

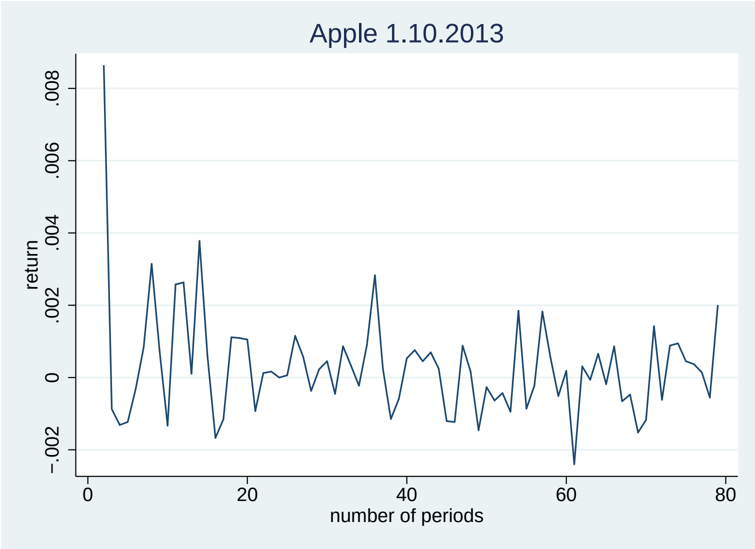
after the opening of the markets. Note that the returns we dropped were those calculated from the first day close price and last previous day close price, so actually, it was rather over night returns than five minute returns.

This phenomenon can also be seen from the line plot 3.2, where we show the feature in particular date 1.10.2013 for Apple Inc. but similar behaviour can be observed for all stocks every trading day.

The first spike represents the over night return and as you can see, it is substantially different from consecutive five minute returns. The reason for such behaviour is a fact that the overnight returns include more information accumulated over the nontrading hours of the night resulting in very different first period's close price of the following day, thus larger absolute return.

The main disadvantage of keeping the over night returns in our analysis is the fact that we would obtained strongly non-stationary time series due to the seasonality of returns which would cause problems while implementing the proposed volatility models. All things considered, we decided to drop them of our analysis which implicitly means that we consider only those news that occured in trading hours, thus avoiding any potential errors. In addition to the first period that we decided to dispose of, we found significant parameters while

Figure 3.1: Apple 5 minute returns plot 1.10.2013



Source: author's computations.

checking for further seasonality in the beginning of the day. Particularly we saw significant relationship between returns and first 15 minutes of the trading day (including the dropped first period) for all the stocks except Google. The signs of the estimation parameters differed both across companies and across time periods. The possible explanation of such behaviour might be the fact that it takes some time for the markets to adjust the price to the informations that arrived through non-trading hours.

We will not drop these observations from the regression for two reasons, it contains informations arrived through the market hours and the behaviour is different for each of the stocks. The solution as we propose it would be to controll for this effects in the mean equation of the GARCH model. In the next section, we will take closer look on the trading volume series, examining the seasonality patterns and introducing the method by which we split the trading volume into expected and unexpected part.

3.1.3 Volume series

As we mentioned in the theoretical framework, we will include both expected and unexpected trading volume to our variance equation of the GARCH model. The volume series, as we utilize it in our analysis, will be modeled as follows:

$$vol_t = \mu + \sum_{i=1}^{p} \beta_i Dperiod_i + \sum_{i=j}^{q} \alpha_j vol_{t-j} \delta t$$

where μ is a constant, $Dperiod_i$ is dummy variable capturing particular period of the day, namely the beginning and the ending periods, vol_{t-j} is just simply the lagged volume value. On the top of that, we check for possible time trends in our regression. The fitted values from this regression would be used as an expected volume, the residuals would be used for unexpected volume which will be further decomposed according to its sign. The summary statistics that we provide below is already filtered from the first period of the day.

Table 3.4: 5 minute volume summary statistic

Company	Mean	Std. Dev.	Min	Max
Apple	29162.35	72022.05	1000	2199099
Exxon Mobil	24958.87	33626.92	1525	634035
Berkshire Hathaway	8826.468	13055.95	315	559321
Microsoft	138238.6	420553.7	28.5	8644737
Google	5752.002	14904.7	100	316600

Source: author's computations.

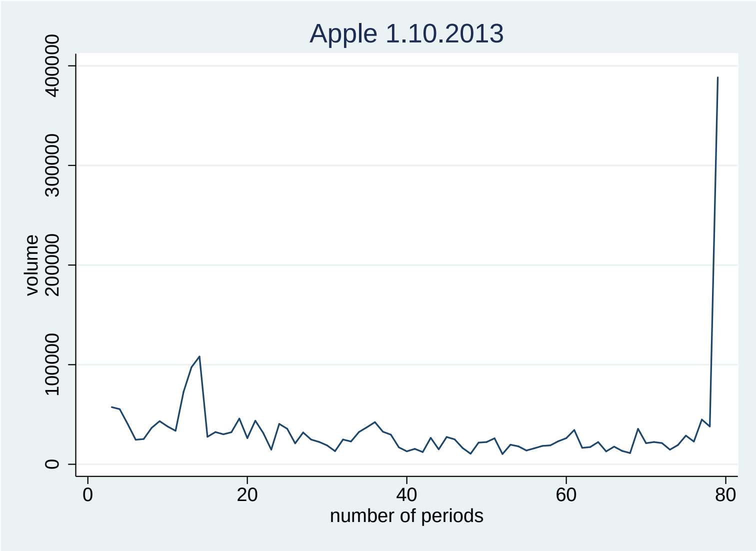
In line with return series, we have observed significantly higher values of volume in the begginning and even more at the end of the trading day. The time span of such behaviour differs across stocks and will be treated in the volume model proposed above.

Generally, it can be seen that the volume series shows J-shaped or U-shaped patterns during the trading day, meaning that the highest values of volume are observed in the first and in the last ten minute period of the day. As an example, let us present you the line plot capturing volume in 1.10.2013. see figure 3.2. This kind of shape of intraday volume is not unique and is observable for all stocks in our analysis. Possible explanation of such volume behaviour is already submited in Admati (1988).

One of the proposed explanation is that high volume in particular time span reveals the presence of asymmetric indormation as noise traders camouflage the activity of the informed traders, thus creates the patterns in volume and volatility. Another explanation enshrined in fundamentals of the technical analysis is that holding security over night represents huge risk, thus daily traders open the possitions in the open and close in the close of the market every day.

Having presented the data we can move further to the model application,

Figure 3.2: Apple 5 minute volume plot 1.10.2013



Source: author's computations.

where we also show the modeling procedure for the Apple Incorporated. For other stocks, the approach will be much the same.

3.2 Model Application

In this section, we move on to actual aplication of our models on the data. In order to make it briefer, we provide the full step by step analysis only for the Apple Inc. For the remaining stocks, the modelling approach is much the same, thus we only show the final estimation results and provide its interpretations. In the end, we will summarize the estimation results of our analysis in the summary.

3.2.1 Apple

Before we start with the modeling of the volatility itself, we should check the assumptions that the model stands on, namely the stacionarity of time series, serial correlations and presence of the ARCH effects in our returns.

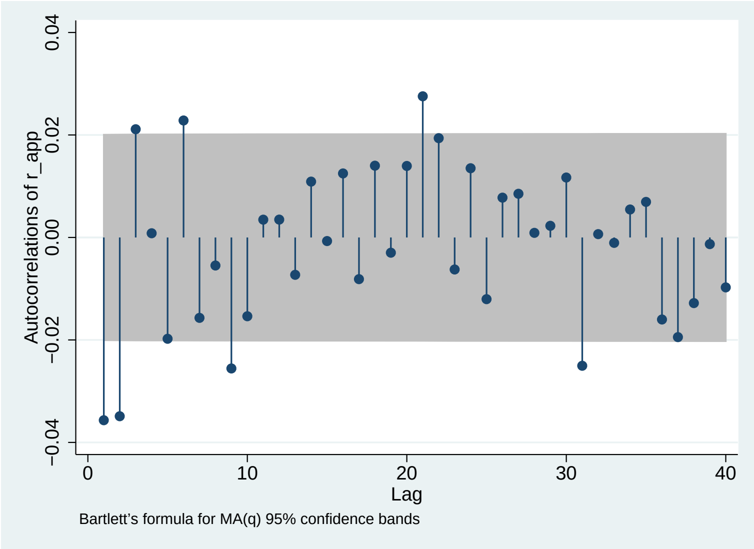
Let us begin with the test of stacionarity assumption, where we use the test proposed by David Dickey and Wayne Fuller so called Dickey-Fuller test for a unit root process. The null hypothessis states that there is unit root process in the series, in other words, the time series is not stacionary. The alternative hypothesis is that we have stacionary series, so we wish to reject the null. Performing the test, we obtain the test statistic equal to -101.227 with the critical value at five percent significance level equal to -2.86. The decision rule connected with this test is that we reject the null hypothesis if our test statistics is less then the critical value, in our case, we can strongly reject the nonstacionarity of the series.

For specification of the mean equation together with controlling the uncorrelation of the returns, we decided to check the autocorrelation function.

We can observe seven significant autocorrelations, at the first, the second and the sixth lag later followed by further lags, namely 21 and 31. This structure of returns can suggest us to use the AR process when identifying the mean equation, above that, we decided to check for the further seasonality as mentioned in the section 3.1.2.

For the mean equation, we decided to run simple AR model. Although the ACF suggests us to include furhter lags, the proper way as we see it is just to include the first three lags, thus obtaining the AR(3) model augmented on the day periods. We have several reasons for this specification, mainly the fact that the inclusion of the additional lag parameters does not contribute much as for the magnitude of their parameters and also does not increase the explanatory

Figure 3.3: ACF of apple 5 minute returns



Source: author's computations.

power of our model. It is also important to note that the joint MLE estimation of our mean and variance equation allows for the heteroscedasticity in the residuals which can can result in insignificance of any of the AR parameter in the mean equation. In such case, we reestimate the mean equation where we drop the insignificant lag parameter. The estimation results of the proposed mean equation are presented in table 3.5.

Table 3.5: Apple mean equation results

Variable	Coefficient	Std. Err.	t	p-value
$L1.r\_app$	-.037061	0.0070786	-5.24	0.000
$L2.r\_app$	-0.0330483	0.0065254	-5.06	0.000
$L3.r\_app$	0.0203862	0.0078847	2.59	0.010
period1	0.0002405	0.0000472	5.10	0.000
period77	0.0004037	0.0001027	3.93	0.000

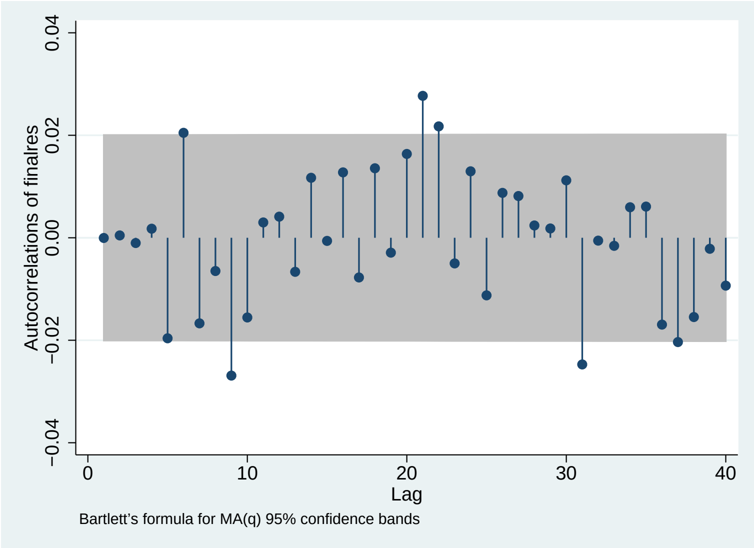
Source: author's computations.

As you can see, we obtained significant coefficient for all explanatory variables. The variables noted as $L1.r\_app$, $L2.r\_app$ until $L3.r\_app$ are lagged returns. $Period1$ stands for the time span (15:35-15:40) and $period77$ stands for the last day's period. As you can see the p-values are close to zero, therefore, the following model is proposed as the mean equation.

$$r_t = \mu + \beta_1 L1.r\_app + \beta_2 L2.r\_app + \beta_3 L3.r\_app + \beta_4 period1 + \beta_5 period77$$

After all, to show you the results of removing the serial correlation from our returns, we present you the ACF of the residuals from the proposed mean equation.

Figure 3.4: ACF Apple mean equation residuals



Source: author's computations.

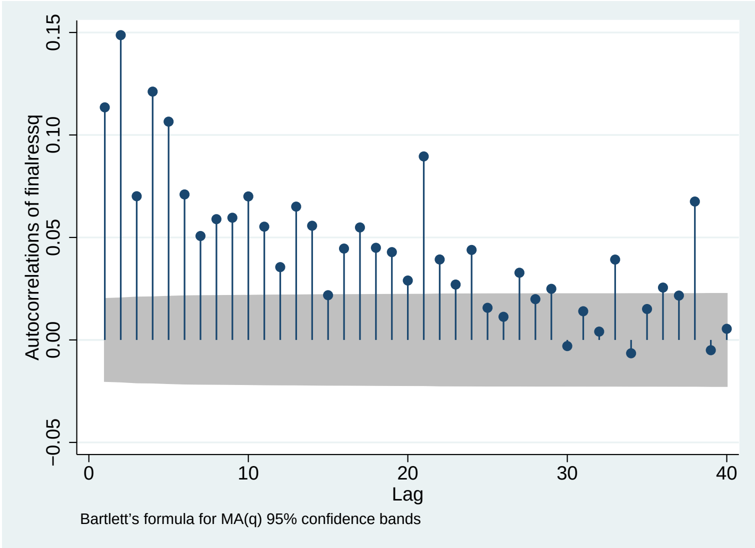
As you can notice, we succeded in removing all major serial correlations, thus the mean equation of our volatility model should be properly specified and we can move to the next task of our analysis which is the testing for the ARCH effects in the residuals.

When testing for the ARCH effects, we begin with the ACF of the squared residuals a_t^2 from our mean equation finding strong dependence up to almost 25th lag.

To confirm the results obtained from the ACF above, we perform the Lagrange-multiplier test for the ARCH effects which is simply equivalent to the usual F statistic for testing joint hypothesis in linear regression of a_t^2 on its lagged values. The results of this test show p-value equal to zero when testing for fitst 15 lags. This results lead us to the conclusion of ARCH presence in residuals.

The last task we want to adress before the estiamtion of the volatility model itself is the proper lag specification of the variance equation. As we mentioned earlier, we will use Akaike Information Criterion when evaluating the lag-length of the model. But first, we need the line to start from. The financial series theory suggest the use of PACF of squared residuals, so first, we will take look at the PACF to identify the starting line and then use AIC for different values

Figure 3.5: ACF Apple mean equation squared residuals



Source: author's computations.

of p,q. After the performance of the AIC test, we found that the best fit is achieved with the (1,1) setting, thus we will estimate the GARCH(1,1) model.

Table 3.6: Apple GARCH(1,1) estimation results

Variable	**Coefficient**	**(Std. Err.)**	**z**	**p-value**
	Mean Equation			
L1.r_app	-0.024	(0.011)	-2.15	0.031
L2.r_app	-0.031	(0.011)	-2.76	0.006
period1	0.000	(0.000)	15.35	0.000
period77	0.000	(0.000)	10.42	0.000
Intercept	0.000	(0.000)	-1.11	0.267
	Variance Equation			
L.arch	0.143	(0.005)	29.77	0.000
L.garch	0.815	(0.005)	153.28	0.000
Intercept	0.000	(0.000)	23.86	0.000

The simple GARCH(1,1) etimation results are provided in the table 3.6. As you can see, all estimation parameters are statistically significant in both equations. As for the variance equation, we can see relatively strong persistance of volatility. The coefficient β equals to 0.815 meaning that previous period's high volatility tend to be followed by current period's high volatility. The responce of the model to the previous shock is not that large according to its magnitude, where α equals to 0.143 but the parameter is very statistically significant. The stacionarity condition of the variance equation is satisfied, since the $\alpha + \beta < 1$. Moreover, both z-scores are large as for the magnitudes. After the inclusion of trading volume, we will observe both the p-values of the GARCH terms and the magnitude of the estimation parameters.

Before we start with the interpretation, let us note that when testing for normality of residuals after the estimation of GARCH model we obtained p-value equal to zero, so the residuals were not normally distributed under the Gaussian distribution in GARCH model, thus we tried different distribution and according the AIC the Student t distribution seems to be the best fit for our model.

Now, let us provide you the table with estimation results of variance equation with both contemporaneous and lagged trading volume terms as described in section above. We will not provide the estimation results of the mean equation, since it is not the area of interest in our analysis.

Table 3.7: Apple GARCH(1,1)-volume estimation results

Variable	**Coefficient**	**(Std. Err.)**	**z**	**p-value**
Contemporaneous Volume Equation				
EVol	-0.0000125	2.98e-06	-4.20	0.000
UvolP	3.00e-06	2.15e-07	13.99	0.000
UVolN	0.0004891	0.0000273	17.93	0.000
L.arch	0.1223221	0.0060873	20.09	0.000
L.garch	0.7352733	0.0080027	91.88	0.000
Lagged Volume Equation				
L.EVol	0.0000173	3.37e-06	5.12	0.000
L.UVolP	3.28e-06	4.60e-07	7.13	0.000
L.UVolN	-0.0000164	7.47e-07	-21.97	0.000
L.arch	0.1432636	0.0050768	28.22	0.000
L.garch	0.780224	0.0063875	122.15	0.000

The interpretation of our regression results is not so straightforward as it may appear, thus we decided to describe each part of the table separately in the following paragraphs.

From the first part of the table which captures the contemporaneous relationship between volume and volatility we can see the different behaviour of the expected volume noted as *Evol* and unexpected volume divided into positive and negative shocks, where *UvolP* represents positive shocks and *UvolN* represents negative shocks. The expected volume is generally considered to represent liquidity trading in financial markets. Since liquidity reduces the riskiness of the market, we expect the parameter to decrease the volatility, as you can see, we obtained negative coefficient that meets our expectations, moreover, this parameter is statistically significant. The parameters capturing the asymmetric effect of volume shocks are also very statistically significant but we expected them to have oposite signs. Generally the unexpected volume represent the information trading which is reflected through volume shock. On the top of that, we expected the positive shock to increase the volatility and negative shock to decrease it. However, our estimation results are showing us that both positive and negative shocks tend to increase the volatility. From the first table, we can also see the change of magnitude of GARCH and ARCH parameters. The GARCH parameter representing the volatility persistance has been reduced almost by 10 percent, after the inclusion of trading volume, and the drop in the z-score is also observable, however, it is stil very statistically significant. The parameter that expresses the response to the previous price shock decreased as well by 14.7 percent. All in all, the results from the contemporaneous regression are in line with the other researches but we can not conclude that the contemporaneous volume is able to explain the persistance of volatility.

As mentioned in the theoretical framework, the approach of including contemporaneous volume was criticised due to assumption of exogeneity and the use of lagged volumes was suggested instead. The results obtained from our regression can show us quite subtle restults. The parameter representing the liquidity trading has positive sign, thus suggesting that expected volume increases volatility in the next period, the magnitude of this parameter is not very large but it is statistically significant. The unexpected volume parameters properly show the presence of asymmetric effect. The positive volume shock can be interpreted that the higher the unexpected volume now the higher the future uncertainty in the markets, thus increases the volatility in the next period. The negative unexpected volume has more of an opposite effect, where both of these parameters are statistically significant. If we focus on the interpretation of our GARCH and ARCH parameters, the drop in magnitudes and z-scores are not as big as in the contemporaneous case, we can observe the drop by 4.2 percent in the GARCH and no decrease in the ARCH term.

In the end, we compared the aproprietness of the three models by the Akaike Information Criterion. We found that both models with included volumes perform significantly better when estimating volatility. Namely the model with lagged volume which carries the predictive ability can be contribution for forcasting future volatility.

3.2.2 Exxon Mobil

The second stock of our analysis is the Exxon Mobil Corporation. We followed the same approach for modeling return, volume and volatility series as in the case of Apple. The resulting model we chose is the GARCH(1,2) later augmented on expected and unexpected volume terms.

The results of our plain GARCH(1,2) regression are provided in the table 3.8. The estimation results of the volume-volatility relationship are in the table 3.9. Before we provide further commentary on the estimation results, we should say that our regression results are in line with our expectations.

Table 3.8: Exxon Mobil GARCH(1,2) estimation results

Variable	**Coefficient**	**Std. Err**	**z**	**p-value**
Mean Equation				
L2.r_exx	-0.0274418	0.0107296	-2.56	0.011
Intercept	0.0000169	7.48e-06	2.26	0.024
Variance Equation				
L1.arch	0.2295221	0.0082482	27.83	0.000
L1.garch	0.3040452	0.0369482	8.23	0.000
L2.garch	0.3641173	0.0335885	10.84	0.000
Intercept	9.12e-08	5.32e-09	17.14	0.000

Table 3.9: Exxon Mobil GARCH(1,1)-volume estimation results

Variable	**Coefficient**	**Std. Err.**	**z**	**p-value**
Contemporaneous Volume Equation				
EVol	9.21e-06	2.09e-06	4.40	0.000
UvolP	0.0000173	9.47e-07	18.32	0.000
UVolN	-0.0000432	1.71e-06	-25.26	0.000
L1.arch	0.1737021	0.0084213	20.63	0.000
L1.garch	0.4344811	0.0535209	8.12	0.000
L2.garch	0.2200761	0.0451006	4.88	0.000
Lagged Volume Equation				
L1.EVol	0.0000151	1.90e-06	7.92	0.000
L1.UVolP	0.000011	2.33e-06	4.71	0.000
L1.UVolN	-0.0000346	3.02e-06	-11.46	0.000
L1.arch	0.1936197	0.007969	24.30	0.000
L1.garch	0.3273208	0.046869	6.98	0.000
L2.garch	0.2889809	0.0402476	7.18	0.000

To comment on the results obtained in our research, we can observe that both in the contemporaneous and in the lagged volume variance equation the expected volume has small positive effect on volatility. More interesting finding is the presence of asymmetric effect in the volume shocks, as you can see, the positive volume shocks tend to increase the volatility and negative volume shocks tend to decrease it. If we focuse on the parameter that represents the persistence of volatility represented by lagged GARCH parameters, we can hardly see substantial differences in their significance, moreover, the sum of the GARCH parameters did not change at all. However, we can observe drop in the ARCH parameters. In the contemporaneous model the ARCH effect was reduced by 22 percent, in the lagged equation the magnitude of change was 19 percent. These observations can be interpreted as there are not substantial decreases of the ARCH and GARCH effects, thus the volume does not explain much. Last thing to note is that all additional variables are very statistically significant. The comparisom of those models by the AIC shows that both contemporaneous and lagged volumes included in the variance equation significantly increase the explanatory power of the model. All in all, we can rather say that the trading volume is positively related to stock return volatility and improves the volatility models but there is not enough evidence to state that inclusion of trading volume explains the GARCH and ARCH parameters.

3.2.3 Berkshire Hathaway

When we examined the Berkshire Hathaway, GARCH(1,1) performed best among tested models. The estimation results are shown in the tables 3.10 and followed by the further commentary.

Table 3.10: Berkshire GARCH(1,1) estimation results

Variable	**Coefficient**	**Std. Err**	**z**	**p-value**
	Mean Equation			
Period2	-0.0002754	0.0000369	-7.46	0.000
Period77	0.0008739	0.0000219	39.87	0.000
	Variance Equation			
L1.arch	0.1746621	0.005994	29.14	0.000
L1.garch	0.7842347	0.0068996	113.66	0.000
Intercept	3.67e-08	2.03e-09	18.10	0.000

Let us begin with the contemporaneous equation. First thing we can notice are the opposite signs of the expected and unexpected volume terms suggesting different reactions of volatility on liquidity trading and information trading. As you can see increase in liquidity trading tend to decrease the volatility in the

Table 3.11: Berkshire GARCH(1,1)-volume estimation results

Variable	**Coefficient**	**Std. Err.**	**z**	**p-value**
Contemporaneous Volume Equation				
EVol	-.0000173	7.66e-06	-2.25	0.024
UvolP	0.0000214	3.62e-07	59.08	0.000
UVolN	0.0012335	0.0000525	23.51	0.000
L1.arch	0.1675081	0.0077001	21.75	0.000
L1.garch	0.6451423	0.0087567	73.67	0.000
Lagged Volume Equation				
L1.EVol	-3.84e-07	7.35e-06	-0.05	0.958
L1.UVolP	0.000013	3.28e-06	3.97	0.000
L1.UVolN	-0.0000873	5.35e-06	-16.31	0.000
L1.arch	0.1785718	0.0073769	24.21	0.000
L1.garch	0.7385407	.0078863	93.65	0.000

market. On the other hand, the arrival of new informations represented by unexpected volumes generally increases the market volatily, however, we can not se asymmetrical effect of unexpected volume shocks on volatility. If we take a look at the volatility parameters, we see quite big decrease both of their magnitude and their z-statistics, however, the parameters are still very statistically significant, thus they should be kept in the model.

In the lagged-volume equation the results are more in line with our expectations. The lagged expected volume has negative sign but it is statisticaly insignificant with the p-value equal to 0.958 which might be interpreted that the lagged expected volume does not affect the future market volatility. The lagged unexpected volume terms nicely capture the asymmetric behaviour of volume shocks. From the estimation results, we can suggest that the positive volume shock (above average volume) tend to increase the future volatility of the market, on the other hand, negative volume shocks tend to decrease it. The drop in magnitude of ARCH and GARCH terms is not as dramatic as in contemporaneous case but it is still present but again both terms still kept their statistical significance, thus they can not be dropped out of our model. When we compare all three models based on our AIC and BIC tests, we find that the volume models perform much better than plain GARCH, thus it can be useful to use the information contained in volume for prediction of future volatility.

To conclude the results, we can say that the estimation is in line with our generall expectations. The expected volume reduces or does not affect the volatility and the asymmetric effect of unexpected volume is present in our analysis. In the next section, we perform the analysis with the Microsoft stocks.

3.2.4 Microsoft

According to AIC we chose the GARCH(1,2) for modeling the conditional variance, above that, the results of Bayesian Information Criterion also confirmed the suitability of the chosen model. The estimation results are provided in table 3.12 and 3.13 .

To comment on the estimation results, we again recieved statistically significant volume variables. In the first case where we examined the effect of contemporaneous volume, we obtained results which are in line with outcomes in previously analyzed stocks. Again, the expected volume has small positive effect on volatility, the positive and negative shocks indicate the pressence of asymmetric effect. Comparing the magnitudes and significance of ARCH and GARCH parameters, we must stay a bit cautious. As you can see, in the contemporaneous volume equation the second lag of conditional variance parameter became statistically insignificant. On the other hand, the first lag increased in the magnitude, thus the persistance given by the sum of magnitudes of GARCH parameters decreased from 0.6903 to 0.6507 which can be represented by 5.74 percentage drop. If we take a look at the ARCH parameter, we can also observe decrease in the magnitude, the percentage change accounts for 22 percent.

Table 3.12: Microsoft GARCH(1,2) estimation results

Variable	**Coefficient**	**Std. Err**	**z**	**p-value**
	Mean Equation			
L1.r_mcsft	-0.0410837	0.0109518	-3.75	0.000
Period2	-0.0001038	0.0000514	-2.02	0.044
Period76	0.0005448	0.0000407	13.39	0.000
Period77	0.0002486	0.0000337	7.37	0.000
	Variance Equation			
L1.arch	0.2578876	0.0085624	30.12	0.000
L1.garch	0.5361835	0.0381386	14.06	0.000
L2.garch	0.1542089	0.0312064	4.94	0.000
Intercept	1.05e-07	5.87e-09	17.94	0.000

The second part of the table 3.13 that captures lagged volume-volatility

Table 3.13: Microsoft GARCH(1,2)-volume estimation results

Variable	Coefficient	Std. Err.	z	p-value
Contemporaneous Volume Equation				
EVol	8.41e-06	7.01e-07	11.99	0.000
UvolP	1.12e-06	7.11e-08	15.76	0.000
UVolN	-3.31e-06	9.77e-08	-33.86	0.000
L1.arch	0.2027905	0.0101869	19.91	0.000
L1.garch	0.6506926	0.0566581	11.48	0.000
L2.garch	0.0314003	0.0440189	0.71	0.476
Lagged Volume Equation				
L1.EVol	7.63e-06	6.24e-07	12.24	0.000
L1.UVolP	1.02e-06	5.77e-08	17.61	0.000
L1.UVolN	-3.20e-06	9.80e-08	-32.63	0.000
L1.arch	0.2382207	0.0095956	24.83	0.000
L1.garch	0.4782216	0.0398169	12.01	0.000
L2.garch	0.160427	0.0321152	5.00	0.000

relationship reports very similar results. All lagged volume parameters are very statistically significant, where the expected and unexpected lagged volumes have similar effects as in the contemporaneous case. The GARCH parameters stayed statistically significant. The decrease of the magnitude can be observed in the first lag of the GARCH parameter but part of it was compensated by the increase of the second lag. The ARCH parameter also decreased but again, stayed very statisticaly significant.

All in all, we can not accept that the inclusion of contemporaneous or lagged volume variables in the variance equation explains the persistance or clustering of the volatility, however, both contemporaneous and lagged volume models perform much better in the terms of model suitability measured by AIC and BIC.

3.2.5 Google

The last stock analyzed in our research is the Google Inc., in the Google case, the return series has not shown any significant autocorrelations, on the top of that, the constant term resulted to be insignificant in the mean equation, thus we provide only the variance equation in the table 3.14. The results after the inclusion of trading volume are shown in the table 3.15 below.

In the model where contemporaneous volume was included, we can observe the same results as in the previous stock, where increased liquidity trading tend to increase the volatility and unexpected volume show asymmetric behaviour. Also the GARCH and ARCH terms were affected like in the Microsoft case. The second GARCH lag became insignificant but the first lag increased in its magnitude. The overall persistance measured by their sum decreased by 8.8 percent, the decrease in the magnitude of the ARCH parameter was 22.3 percent which is very similar to the results obtained in the previous stocks.

Table 3.14: Google GARCH(1,2) estimation results

Variable	Coefficient	Std. Err.	t	p-value
L1.arch	0.2146496	0.0098549	21.78	0.000
L1.garch	0.3882413	0.0464662	8.36	0.000
L2.garch	0.3506986	0.0399216	8.78	0.000
Intercept	6.28e-08	3.20e-09	19.60	0.000

Source: author's computations.

Table 3.15: Google GARCH(1,2)-volume estimation results

Variable	**Coefficient**	**Std. Err.**	**z**	**p-value**
Contemporaneous Volume Equation				
EVol	0.0000457	8.96e-06	5.11	0.000
UvolP	0.000025	4.00e-06	6.25	0.000
UVolN	-0.0000898	5.12e-06	-17.55	0.000
L1.arch	0.166873	0.0110383	15.12	0.000
L1.garch	0.6750936	0.0758509	8.90	0.000
L2.garch	0.0946877	0.0633682	1.49	0.135
Lagged Volume Equation				
L1.EVol	0.0000442	7.77e-06	5.69	0.000
L1.UVolP	0.0000231	3.03e-06	7.64	0.000
L1.UVolN	-0.0000896	4.55e-06	-19.68	0.000
L1.arch	0.201705	0.0103598	19.47	0.000
L1.garch	0.4201113	0.050495	8.32	0.000
L2.garch	0.3004802	0.0422685	7.11	0.000

The estimation results from the equation with lagged volumes also show similar resutls as in the previous case. The lagged parameters are statisticaly significant, where expected volume should increase the volatility in the next period and the volume shock asymmetrically effect the future volatility depending on the sign of the shock. As for the magnitude of the common GARCH and ARCH parameters, the sum of the GARCH terms decreased by 2.48 percent and the ARCH parameter itself decreased by 6 percent. As you can see both stayed statistically significant, thus we hardly have evidence that trading volume is able to explain the volatility.

When comparing the three models by AIC and BIC, both volume models perform significantly better for modeling conditional variance of the stock returns, thus trading volume can provide us useful information for modeling volatility. In the next section, we provide overall summary of our estimation results obtained in our analysis.

3.3 Summary

Firstly, we had to choose appropriate data to conduct our analysis on. Since we were examining the intraday volume-volatility relationship, we had to choose market with high volumes, thus we focused on the two biggest stock exchanges headquartered in New York, the NYSE and NASDAQ from which we picked top five companies with the highest market capitalisation, namely the Apple, Exxon Mobil, Berkshire Hathaway, Microsoft and Google. Our next task was to analyze the data which resulted in quite big data sets with more than nine thousand observations per stock. Before employment of the volume-volatility model itself, we had to test the time series data for stacionarity which shown that both return and volume series are stacionary. The next task was to identify proper structure of the model for modeling mean equation of returns and volume series. The raw log returns were showing minor autocorrelations and seasonality sensitive for the beginning and the end of the trading day, thus simple AR proces augmented on seasonal dummy variables turned out to be suitable for modeling the mean equation. The next step was to test for the presence of ARCH effect in the residuals from the mean equation, where the autocorrelation function indicated strong dependence later confirmed by ARCH-LM test, thus the employment of GARCH model seemed to be appropriate. When identifying the proper lag-length of the GARCH model, we used the Akaike Information Criterion. From the comparison of 25 models for each stock we choose the best one later used in our analysis. Before the inclusion of trading volume in our conditional variance equation, we had to properly decompose the volume into expected and unexpected parts. Generally the fitted values from AR model with seasonal dummy variables for volume series served as the expected volume, the residuals as the unexpected volume, above that, we split the unexpected volume into positive and negative parts in order to capture the asymmetric effect of volume shocks on volatility. Once the pre estimation issues were done, we could finally approach to the modelling of mentioned relationship. First we estimated the plain GARCH(p,q) model followed by the estimation of the model augmented on contemporaneous volume variables and in the end with the lagged volume variables. The estimation results of the plain GARCH model showed strong presence of volatility persistance. After the inclusion of contemporaneous volume we have seen that all the volume parameters are statistically significant which also confirms the approprietness of spiting the volume into expected and unexpected part. In the case of contemporaneous expected volume interpreted as liquidity trading, we could observe the positive relationship, meaning that the increase in liquidity trading tend to increase the stock return volatility. The magnitude of these effects is not that large, where we are speaking about the scale of 10^{-5} up to 10^{-7} . In order to increase the volatility in units, there would have to be approximately additional hundred thousand or more stocks traded in each five minute interval. The estimation results of asymmetric effect of unexpected volume shown that positive volume shocks tend to increase the volatility. This effect is being attributed to the information trading. On the other hand the negative shocks tend to decrease the volatility. The scales of unexpected volume terms are much like in the expected volume. Apart from the information of volume-volatility relationship we wanted to examine the change of the magnitude of persistance parameters. Although some earlier studies documented that the inclusion of volume should significantly reduce the traditional GARCH parameters, our findings did not confirmed this conclusion. However, the presence of both contemporaneous and lagged volume in our volatility equation results in significantly better fit confirmed by both Akaike Information Criterion and Bayesian Information Criterion, especially the inclusion of lagged volume may improve the predictive ability of the GARCH model.

Chapter 4

Conclusion

Although the volume volatility relationship is broadly described by many prior researchers, there are still rather conflicting results that drive the econometricians and financial analysts to deeper considerations of the models and researches with different asset classes or time spans of the data. Generaly there are two common theories explaining the relationship of trading volume and asset return volatility. The theory of information flows that connects the flow of information with the behaviour of the financial markets and the dispersion of beliefs theory which speaks about the different behaviour of the market participants.

In our analysis we decided to examine the Mixture of Distribution Hypothesis which is a part of the theory of information flows. In a nutshell, the theory states that both volume and volatiliy are simultaneously generated by random sequence of informations that flow to the market. Following the approach of Lamoureux (1990) who has augmented the ARCH model on the raw volume series and found that after the inclusion of the trading volume, the ARCH parameters became insignificant. However, in our research we extended the analysis by using more appropriate GARCH model. Moreover, instead of using the raw trading volume, we decided to split the volume into two parts, the expected volume representing the liquidity trading and the unexpected volume representing the information trading. On the top of that, the unexpected volume is further split to capture the different effect of positive and negative volume shocks. As you may suspect, the liquidity trading represented by expected volume is stabilizing factor of the markets, thus we expected it to reduce the stock return volatility. The unexpected volume split into positive and negative shocks is expected to have asymmetric effect on volatility, where positive volume shocks should be positively related to volatility whereas negative shocks should decrease it. Before presenting the estimation results, we should mention that volatility and trading volume are both determined by the inflow of information, where the contemporaneous volume should not be assumed to be exogenous in the model, thus we performed also the regression with lagged trading volumes.

For the purpose of our analysis, we decided to choose top five American stocks that are being traded in the NYSE or NASDAQ. The stocks are constituents of S&P500 market index and represent the top five companies in the terms of the market capitalisation. For each stock, we have collected sufficient amount of data on the five minute intraday basis roughly for six consecutive months, thus we did not have to worry about problems that occure with small data sets. On the top of that, all datasets we collected behaved properly in the sence of stationarity and other desired statistical properties.

The estimation results of both the contemporaneous and lagged models showed us few interesting facts. The general message we should take from this research is that the trading volume is positively related to volatility, meaning that the increase in trading volume tends to increase the conditional variance of returns. The volume parameters representing different types of trading have significant effects on volatility, thus the split of the volume into expected and unexpected terms was justified. Despite the fact that we expected the liquidity trading to have stabilising effect on volatility, we obtained quite opposite results. The expected volume parameter has positive, statistically significant effect. The unexpected volume divided into positive and negative parts showed the presence of asymmetric effect, where the positive volume shocks tend to increase the volatility and negative volume shocks tend to decrease it. This findings generally meet our expectations. When looking at the GARCH and ARCH parameters, we are not able to confirm that either lagged or contemporaneous volumes reduce their significance. For all stock after the inclusion of trading volumes the persistance parameters stayed significant but the slight reduction of their magnitude was observable, thus we can conclude that the inclusion of trading volume changes the magnitude of GARCH and ARCH parameters but does not affect their significance. In case of Microsoft and Google, the inclusion of contemporaneous trading volume made the second GARCH parameter statistically insignificant, however the effect was transfered to the first lag of GARCH parameter thus the overall GARCH effect was reduced just slightly.

The comparison of the plain GARCH models with those where trading volume was included showed us that the volume models perform significantly better. This can be benefiting for the prediction of the future volatility. In the end, we would like to adress few things for future research.

Focusing more on the significance of ARCH and GARCH parameters, employment of different kind of modeling approaches could also show some new results in our topic and contribute to further analysis of the problem. It is good to note that each stock in our analysis behaved a bit differently, thus instead of analysing the individual stocks, it would be interesting to aggregate the data and use the stock market indices. This generalisation would provide us comparison tool for tracking distinctions of the behaviour of the different stock markets around the world. As you may know, in the period we used in our analysis nothing extraordinary happened in terms of global economy, thus the analysis is conducted on the data in common market situation. For the next research, it could also be interesting to invetigate whether there are some substantial differences in the volume volatility relationship in times of global economy crisis, where stock markets experiance downturn together with substantionaly higher volatility.

Bibliography

ADMATI, P. (1988): "Theory of Intraday Patterns: Volume and price Variability." *The Review of Financial Studies* **1(1)**: pp. 3–40.

AKAIKE, H. (1973): "Information theory as an extension of the maximum likelihood principle." *In B. N. Petrov and F. Csaki (Eds.), Second International Symposium on Information Theory, Akademiai Kiado, Budapest* pp. 267–281.

ALI F. DARRAT, Shafiqur Rahman, M. Z. (2003): "Intraday trading volume and return volatility of the DJIA stocks: A note." *Journal of Banking and Finance* **(27)**: pp. 235–243.

BLACK, F. (1976): "Studies of Stock Price Volatility Changes." *Proceedings of the Business and Economics Section of the American Statistical Association* pp. 177–181.

BOLLERSLEV, T. (1986): "Generalizet autoregresive conditional heteroscedastic." *Journal of Econometrics* **31**: pp. 307–327.

CHEN, Firth, R. (2001): "The dynamic relation between stock returns, trading volume, and volatility." *The financial review* **38**: pp. 153–174.

CLARK, P. K. (1973): "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." *Econometrica* **41(1)**: pp. 135–155.

COPELAND, T. E. (1976): "A MODEL OF ASSET TRADING UNDER THE ASSUMPTION OF SEQUENTIAL INFORMATION ARRIVAL." *The Journal of Finance* **31(4)**: pp. 1149–1168.

ENGLE, R. (1982): "Autoregresive conditional heteroscedasticity with estimates of the variance of the United Kingdom inflation." *Econometrica* **50(4)**: pp. 987–1007.

KENNETH P. BURNHAM, D. R. A. (2002): *Model Selection and Multimodel Inference*. Washington, DC: FIAS.

LAMOUREUX, L. (1990): "Heteroskedasticity in Stock Return Data: Volume versus GARCH effects." *The Journal of Finance* **45(1)**: pp. 221–229.

LAMOUREUX, L. (1994): "Endogenous Trading Volume and Momentum in Stock-Return Volatility." *Journal of Business and Economic Statistics* **12**: pp. 253–260.

L.R. GLOSTEN, R. Jagannthan, D. R. (1993): "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *The Journal of Finance* **48(5)**: pp. 1779–1801.

MANDELBROT, B. (1963): "The variation of certain speculative prices." *Journal of Business* **15(3)**: pp. 131–150.

NAJAND, Y. (1991): "Examination of the Relationship Between Volume and Price Variability in Futures Markets." *Journal of Futures Markets* **11**: pp. 613–621.

NELSON, D. (1991): "Conditional heteroscedasticity in asset returns: A new approach." *Econometrica* **59(2)**: pp. 347–370.

RUEY S. TSAY, W. (2005): *Analysis of Financial Series*. Chicago: John Wiley & Sons, Inc.

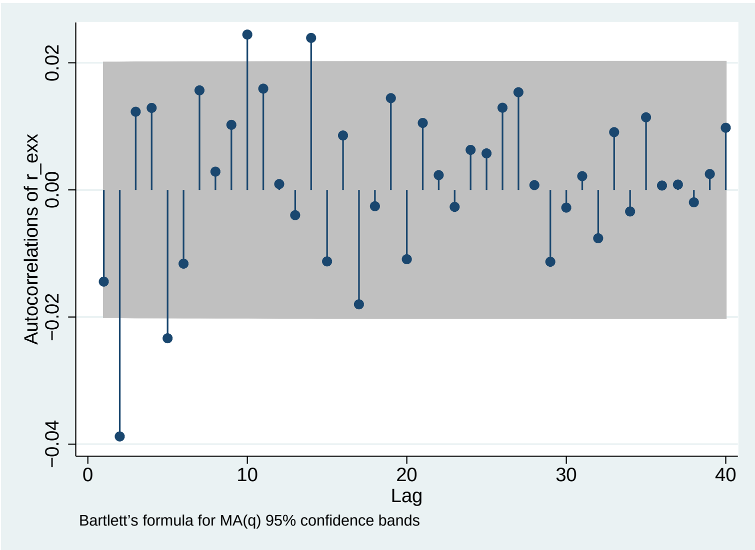
S. KULLBLACK, R. A. L. (1951): "On information sufficiency." *Annals of Mathematical Statistics* **22(1)**: pp. 79–86.

ZAKOIAN, J. (1994): "Threshold heteroscedastic functions." *Journal of Economic Dynamics and Control* **18(5)**: pp. 931–955.

Appendix A

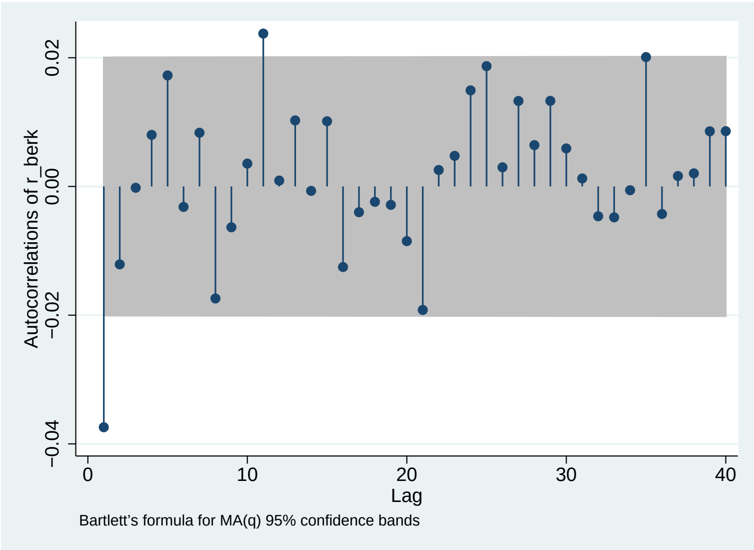
Figures

Figure A.1: ACF of Exxon 5 minute returns



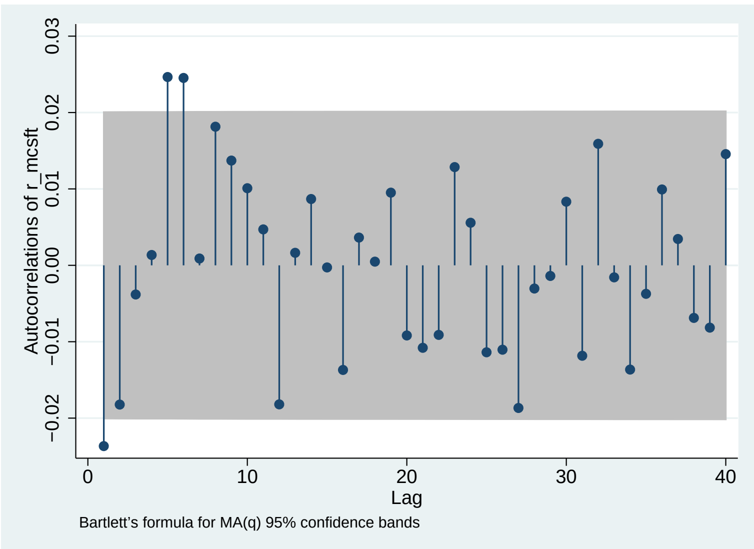
Source: author's computations.

Figure A.2: ACF of Berkshire 5 minute returns



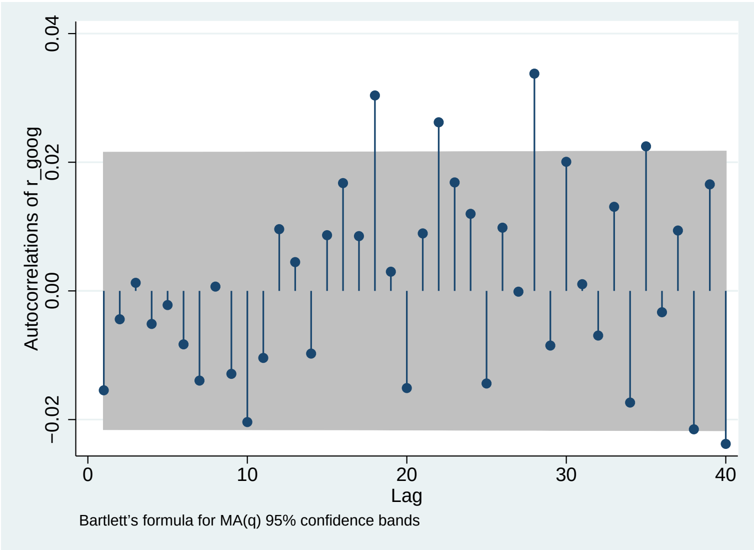
Source: author's computations.

Figure A.3: ACF of Microsoft 5 minute returns



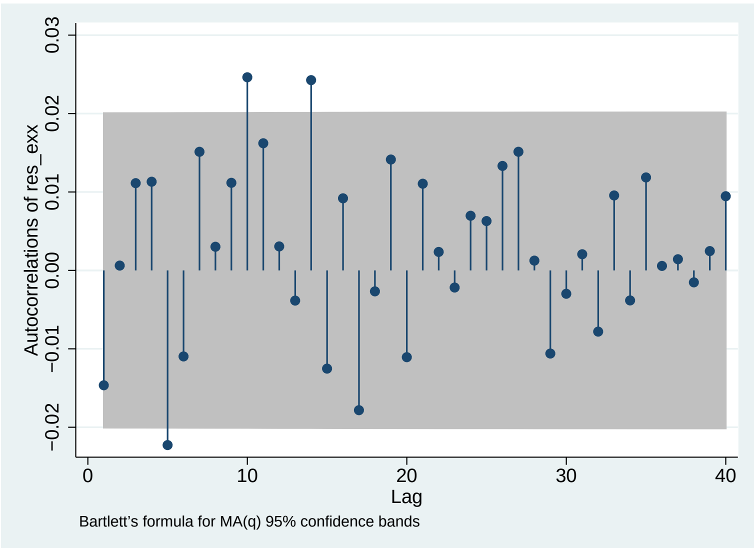
Source: author's computations.

Figure A.4: ACF of Google 5 minute returns



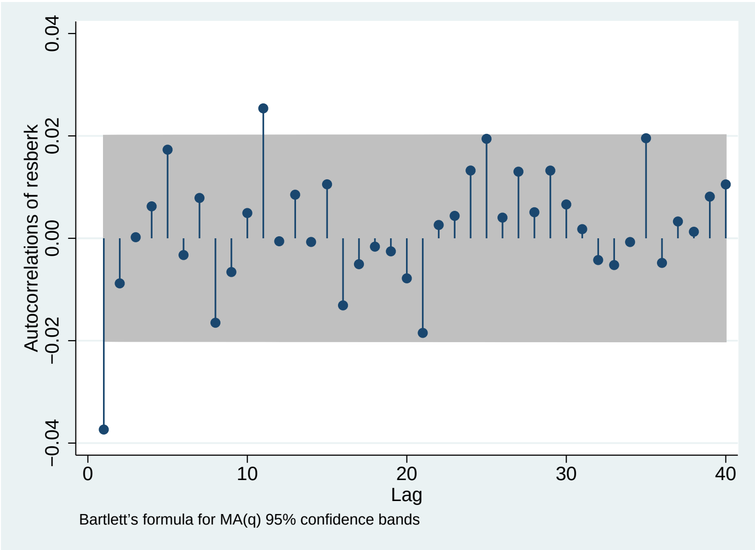
Source: author's computations.

Figure A.5: ACF of Exxon 5 minute return residuals



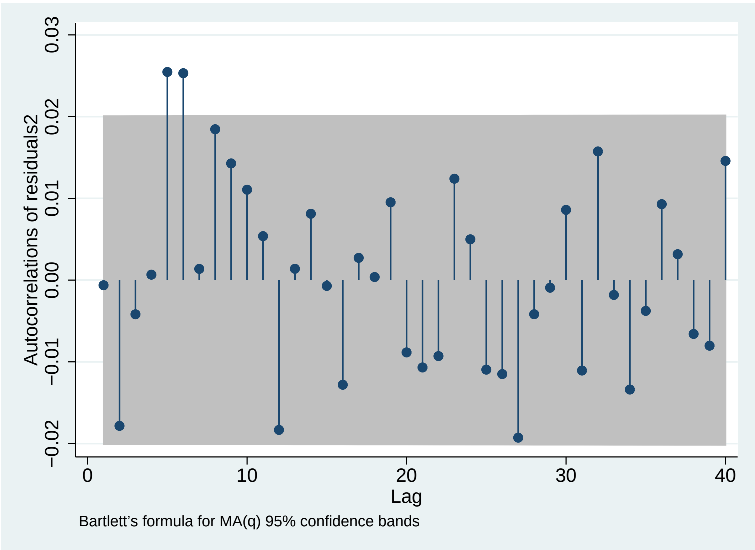
Source: author's computations.

Figure A.6: ACF of Berkshire 5 minute return residuals



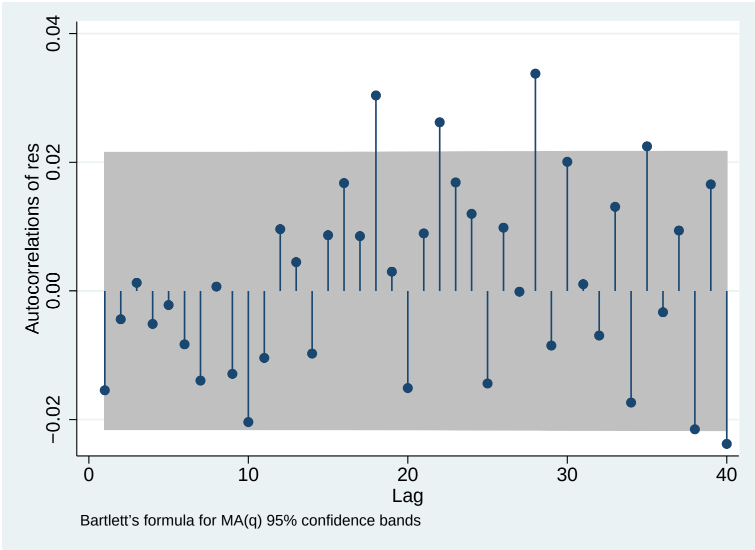
Source: author's computations.

Figure A.7: ACF of Microsoft 5 minute return residuals



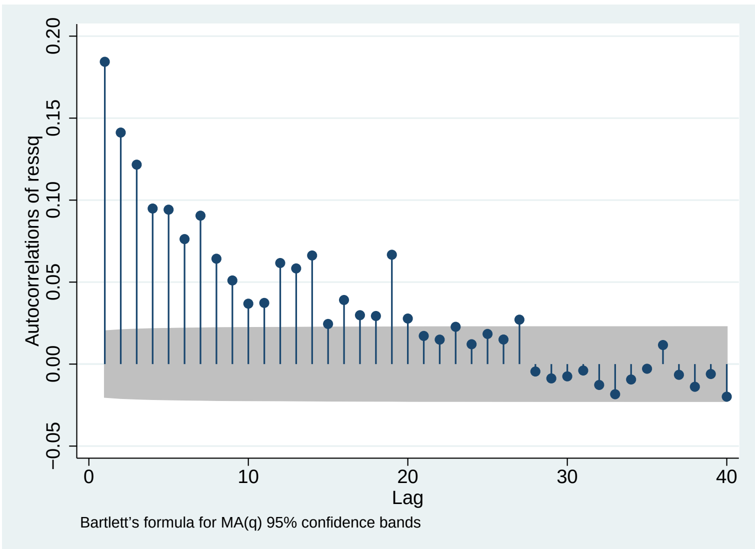
Source: author's computations.

Figure A.8: ACF of Google 5 minute return residuals



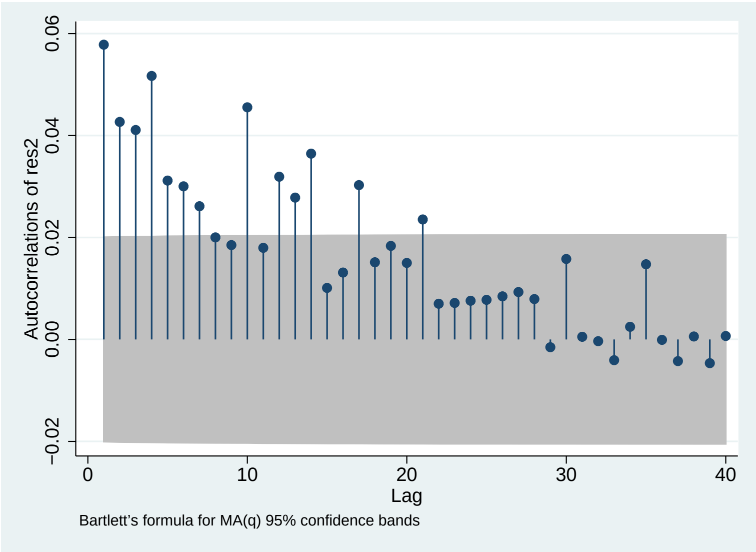
Source: author's computations.

Figure A.9: ACF of Exxon squared residuals



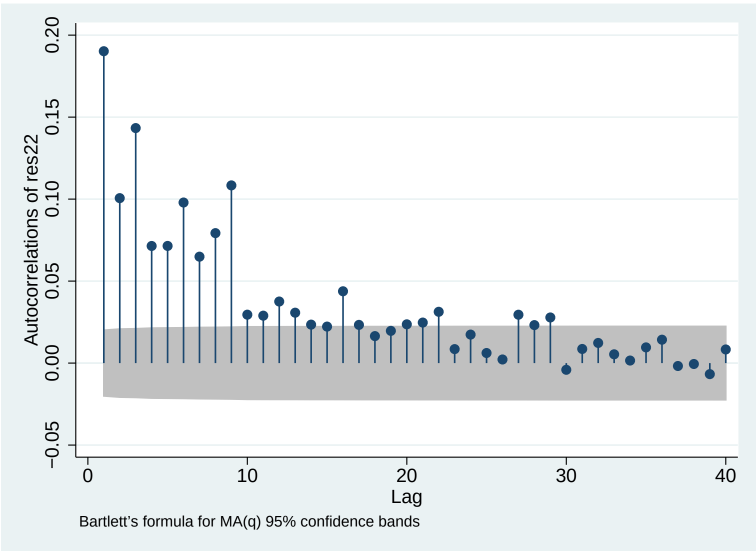
Source: author's computations.

Figure A.10: ACF of Berkshire squared residuals



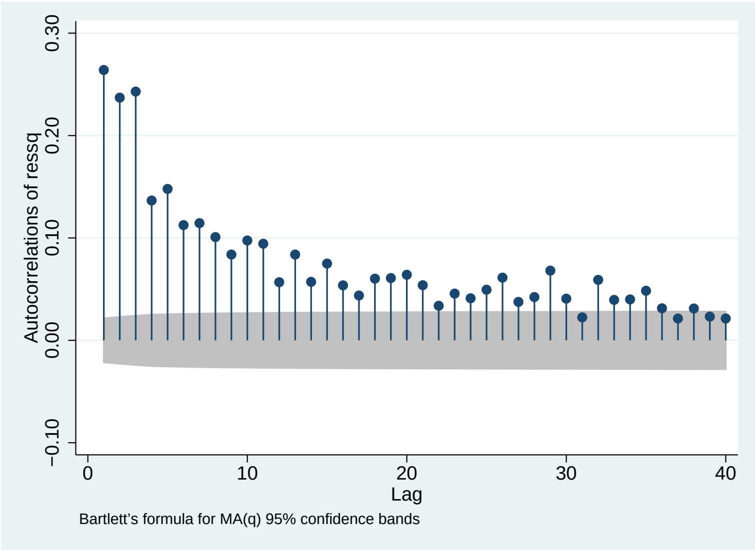
Source: author's computations.

Figure A.11: ACF of Microsoft squared residuals



Source: author's computations.

Figure A.12: ACF of Google squared residuals



Source: author's computations.

Figure A.13: ARCH-LM Apple

lags(p)	chi2	df	Prob > chi2
1	123.814	1	0.0000
2	302.037	2	0.0000
3	318.236	3	0.0000
4	401.510	4	0.0000
5	453.244	5	0.0000
6	459.390	6	0.0000
7	460.248	7	0.0000
8	464.792	8	0.0000
9	470.842	9	0.0000
10	482.362	10	0.0000
11	487.285	11	0.0000
12	487.328	12	0.0000
13	499.605	13	0.0000
14	504.991	14	0.0000
15	507.211	15	0.0000

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Source: author's computations.

Figure A.14: ARCH-LM Exxon

lags(p)	chi2	df	Prob > chi2
1	325.756	1	0.0000
2	438.057	2	0.0000
3	499.216	3	0.0000
4	521.508	4	0.0000
5	545.889	5	0.0000
6	554.922	6	0.0000
7	577.800	7	0.0000
8	581.034	8	0.0000
9	581.929	9	0.0000
10	581.872	10	0.0000
11	582.374	11	0.0000
12	594.087	12	0.0000
13	600.770	13	0.0000
14	610.705	14	0.0000
15	612.657	15	0.0000

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Source: author's computations.

Figure A.15: ARCH-LM Berkshire

lags(p)	chi2	df	Prob > chi2
1	31.943	1	0.0000
2	46.744	2	0.0000
3	59.551	3	0.0000
4	79.681	4	0.0000
5	84.716	5	0.0000
6	89.418	6	0.0000
7	92.483	7	0.0000
8	93.787	8	0.0000
9	94.935	9	0.0000
10	109.048	10	0.0000
11	109.724	11	0.0000
12	114.784	12	0.0000
13	118.049	13	0.0000
14	124.606	14	0.0000
15	124.556	15	0.0000

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Source: author's computations.

Figure A.16: ARCH-LM Microsoft

lags(p)	chi2	df	Prob > chi2
1	345.557	1	0.0000
2	386.720	2	0.0000
3	513.215	3	0.0000
4	517.312	4	0.0000
5	531.553	5	0.0000
6	565.612	6	0.0000
7	570.550	7	0.0000
8	588.702	8	0.0000
9	628.398	9	0.0000
10	632.822	10	0.0000
11	633.282	11	0.0000
12	633.410	12	0.0000
13	634.128	13	0.0000
14	634.121	14	0.0000
15	634.056	15	0.0000

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Source: author's computations.

Figure A.17: ARCH-LM Google

lags(p)	chi2	df	Prob > chi2
1	580.555	1	0.0000
2	831.978	2	0.0000
3	1024.218	3	0.0000
4	1026.449	4	0.0000
5	1046.629	5	0.0000
6	1048.784	6	0.0000
7	1061.511	7	0.0000
8	1064.989	8	0.0000
9	1066.295	9	0.0000
10	1074.126	10	0.0000
11	1081.394	11	0.0000
12	1082.550	12	0.0000
13	1088.597	13	0.0000
14	1088.565	14	0.0000
15	1095.399	15	0.0000

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Source: author's computations.

Appendix B

Tables

Table B.1: Dickey-Fuller test for returns

Stock	Statistic value	Critical value	p-value
Apple	-101.227	-2.860	0
Exxon Mobil	-99.158	-2.860	0
Berkshire Hathaway	-101.468	-2.860	0
Microsoft	-100.081	-2.860	0
Google	-92.663	-2.860	0

Source: author's computations.

Table B.2: Dickey-Fuller test for volume

Stock	Statistic value	Critical value	p-value
Apple	-77.882	-2.860	0
Exxon Mobil	-62.765	-62.765	0
Berkshire Hathaway	-72.579	-2.860	0
Microsoft	-87.489	-2.860	0
Google	-75.137	-2.860	0

Source: author's computations.

Table B.3: Akaike information criterion Apple

p/q	1	2	3	4	5
1	-104879.5	-104876.1	-104875.2	-104875.4	-104876.3
2	-104874.9	-104875.6	-104878.51	-104876.3	-104875.5
3	-104877.1	-104878.9	-104877.3	-104879.1	-104878.51
4	-104876.8	-104876.3	-104876.35	-104878.3	-104876.9
5	-104877.54	-104877.78	-104875.35	-104875.9	-104875.1

Source: author's computations.

Table B.4: Akaike information criterion Exxon

p/q	1	2	3	4	5
1	-108279.3	-108316.1	-108315.2	-108315.2	-108314.3
2	-108314.9	-108315.6	-108314.51	-108314.3	-108313.5
3	-108315.8	-108314.92	-108313.7	-108313.6	-108314.41
4	-108313.8	-108312.9	-108313.21	-108315.3	-108314.9
5	-108314.26	-108315.33	-108315.7	-108315.6	-108315.7

Source: author's computations.

Table B.5: Akaike information criterion Berkshire

p/q	1	2	3	4	5
1	-110880.6	-110878.7	-110877.1	-110879.9	-110879.8
2	-110878.7	-110878.3	-110879	-110879.2	-110878.6
3	-110877.2	-110878.1	-110878.7	-110878.6	-110879.1
4	-110880.3	-110878.6	-110879.1	-110878.9	-110879.3
5	-110879.6	-110880	-110879.7	-110879.4	-110878.9

Source: author's computations.

Table B.6: Akaike information criterion Microsoft

p/q	1	2	3	4	5
1	-103694.2	-103698.9	-103698.3	-103697.4	-103697.8
2	-103698.2	-103697.4	-103697.7	-103697.1	-103696.6
3	-103697	-103698.6	-103698.4	-103698.1	-103697.7
4	-103716.3	-103715.3	-103714.2	-103712.4	-103712
5	-103715.2	-103715	-103711.2	-103712.9	-103712

Source: author's computations.

Table B.7: Akaike information criterion Google

p/q	1	2	3	4	5
1	-92529.65	-92549.73	-92547.16	-92547.2	-92548.1
2	-92547.6	-92548.3	-92548.2	-92547.1	-92548.12
3	-92548	-92549.3	-92548.9	-92547.3	-92547
4	-92548.4	-92548.3	-92549.1	-92547.3	-92548.61
5	-92547.9	-92549.3	-92549	-92548.96	-92549.66

Source: author's computations.

Table B.8: Post estimation AIC

Stock	AIC-GARCH	AIC-Volume-GARCH	AIC-L.Volume-GARCH
Apple	-104879.5	-105883.5	-105655.3
Exxon	-108316.1	-108718	-1086572
Berkshire	-110880.6	-111716.3	-110995.6
Microsoft	-103698.9	-104667	-104658
Google	-92549.3	-93078.5	-92856

Source: author's computations.

Table B.9: Apple cross distribution comparison

Distribution	AIC-GARCH	AIC-Volume-GARCH	AIC-L.Volume-GARCH
Student	-104879.5	-105883.5	-105655.3
Gaussian	-103972.7	-104128.87	-104088.41

Source: author's computations.

Table B.10: Exxon cross distribution comparison

Distribution	AIC-GARCH	AIC-Volume-GARCH	AIC-L.Volume-GARCH
Student	-108316.1	-108718	-1086572
Gaussian	-107942.3	-108236.1	-108212.4

Source: author's computations.

Table B.11: Berkshire cross distribution comparison

Distribution	AIC-GARCH	AIC-Volume-GARCH	AIC-L.Volume-GARCH
Student	-110880.6	-111716.3	-110995.6
Gaussian	-110502.9	-111198.7	-110654.5

Source: author's computations.

Table B.12: Microsoft cross distribution comparison

Distribution	AIC-GARCH	AIC-Volume-GARCH	AIC-L.Volume-GARCH
Student	-103698.9	-104667	-104658
Gaussian	-103426	-104419.9	-104432.5

Source: author's computations.

Table B.13: Google cross distribution comparison

Distribution	AIC-GARCH	AIC-Volume-GARCH	AIC-L.Volume-GARCH
Student	-92549.3	-93078.5	-92856
Gaussian	-92147	-92789	-92527

Source: author's computations.