Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Understanding systematic risk of assets at various quantiles of return distribution

Author: **Tomáš Rusý**

Supervisor: **PhDr. Jozef Baruník, Ph.D.**

Academic Year: **2015/2016**

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, January 3, 2016

Signature

Acknowledgments

At this place, I would like first to acknowledge my supervisor PhDr. Jozef Baruník, Ph.D., who turned up with advice when I needed it and who found time for me even when no one would expect him to do so.

Secondly, I would like to thank to Dr. Ian Wood, who provided me with some pieces of advice regarding my bachelor thesis during my time in Australia, despite being extremely busy at that time.

Last, but certainly not least, I would like to thank to my lovely Eliška, whose endless support and care have helped me to keep my focus and energy levels high which was necessary to complete this thesis.

Thank you.

Abstract

In this thesis, we deal with the application of quantile regression to the Capital Asset Pricing Model, which is derived in the thesis. We investigate a real dataset to determine if one of many implications – constant beta at different quantiles of return distribution, of the model is met. For that purpose, we use Khmaladze test which is perfectly suited for testing if asset's beta varies over return distribution. Before we run the test we introduce both quantile regression and the Khmaladze test to the reader in simple and clear notation as we do not expect the reader to be familiar with this regression technique.

JEL Classification	C50, C51, C58, C13
Keywords	Capital Asset Pricing Model, Quantile Regression, Khmaladze Test

Author's e-mail	rusy.tomas@seznam.cz
Supervisor's e-mail	barunik@fsv.cuni.cz

Abstrakt

V této práci se zabýváme analýzou modelu oceňování kapitálových aktiv, který je v práci odvozen, pomocí kvantilové regrese. Analýza je provedena na reálných datech, na kterých zkoumáme, zda–li je splněn jeden z mnoha důsledků modelu, a to že je beta konstantní v různých kvantilech distribuční funkce výnosu. K tomu nám poslouží Khmaladzeho test, který se pro testování měnící se bety v kvantilech distribuční funkce výnosu perfektně hodí. Jak kvantilovou regresi, tak Khmaladzeho test navíc před samotným testováním v jednoduchém a přehledném značení uvedeme a vysvětlíme, tudíž jejich znalost k porozumění práce není nutná.

Klasifikace JEL	C50, C51, C58, C13
Klíčová slova	Model oceňování kapitálových aktiv, Kvantilová regrese, Khmaladzeho Test

E-mail autora	rusy.tomas@seznam.cz
E-mail vedoucího práce	barunik@fsv.cuni.cz

Contents

List of Tables

List of Figures

Acronyms

OLS Ordinary Least Squares

CAPM Capital Asset Pricing Model

MVF Minimum Variance Frontier

CML Capital Market Line

SML Security Market Line

Bachelor Thesis Proposal

Author	Tomáš Rusý
Supervisor	PhDr. Jozef Baruník, Ph.D.
Proposed topic	Understanding systematic risk of assets at various quantiles of return distribution

Topic characteristics The year was 1964 when William Sharpe introduced the capital asset pricing model (CAPM). It was the first ever piece in asset pricing theory, yet still after more than 50 years it is widely used in applications. However, there appear to be few difficulties in practical use. For example an assumption about constant beta over return distributions proves to be limiting for effective use of common ordinary least squares regression. For this purpose, we introduce quantile regression, which in this case shows much better results.

Hypotheses At the very beginning of my work regarding my bachelor thesis, I put together a few questions which I would like to have been answered once the thesis is submitted. Firstly, the CAPM assumes linear relationship between the expected return of a stock and its risk. This hypothesis will be tested against a surmise that risk return relation varies in different quantiles of a return distribution. Finally, I would like to run quantile regression on relevant data in order to obtain evidence about varying beta, a thing which is not captured by ordinary least squares regression.

Methodology The thesis will use ordinary least squares regression to analyze the effectiveness of the CAPM. Moreover, quantile regression will be presented as a technique to describe varying beta over return distribution.

Outline

1. Introduction
2. Establishing the CAPM
3. Difficulties of the CAPM
4. Ordinary Least Squares
5. Estimation of the CAPM Using Quantile Regression
6. Quantile Regression, Analysis of the Data
7. Conclusion

Core bibliography

1. Fama, E. & French K. (2004): "capital asset pricing model: theory and evidence." *Journal of Economic Perspectives* **18**: pp. 25–46.
2. Koenker, R. & Bassett G. (1978): "Regression quantiles." *Econometrica*, Econometric Society **46(1)**: pp. 33–50.
3. Koenker, R. & Hallock K. (2001): "Quantile regression." *Journal of Economic Perspectives* **15(4)**: pp. 143–156.
4. Sharpe, W. (1964): "Capital asset prices: a theory of market equilibrium under conditions of risk." *Journal of Finance* **19(3)**: pp. 425–442.

Author Supervisor

Chapter 1

Introduction

This bachelor thesis deals with a model in the asset pricing theory, with the Capital Asset Pricing Model (CAPM). Originally established by Sharpe (1964) and Lintner (1965), it became one of the most used models in investment portfolio analysis, mainly because of its simplicity. The model states, that the return of an asset, which is of interest, is affected only by market return, which is the only variable which we can use to predict the systematic risk of an asset return. It moreover states, that the effect of market return on the asset return is linear. This would all seem to be rather a naive – or in more formal words too simple model. Therefore it comes as a no surprise, that number of people have already showed, by whole panoply of ways, that the model fails empirically, most notably Fama & French (2004). In this thesis, we will follow an approach of number of econometricians, who tested an implication of the CAPM – that market beta stays same in all quantiles of return distribution. This can be analysed by quantile regression of Koenker & Bassett (1978) as was done for example by Barnes & Hughes (2002). They analysed asset's beta for over–performing and under–performing firms and found for example that a size of a firm plays a significant role when determining asset return for an under–performing firm. Another consequence of the CAPM was tested by Chang *et al.* (2011) as they found out, that not in all quantiles is the relationship captured by the CAPM positive. Allen *et al.* (2009) went even further with application of quantile regression to the CAPM and analysed the three factor model presented by Fama & French (2004).

We will focus on the original CAPM and our main objective will be to analyse if asset's betas change in different quantiles of return distribution. As we

stated the implication of the CAPM is that they do not change, however, there are plenty of economic reasons why we could expect that some companies are rewarded more when over–performing or on the other hand are losing more when under–performing. Why we should expect this to be so is relatively simple. Imagine an industry and a small and a big company in that industry. One could expect that when a big firm is under–performing (getting smaller returns than the CAPM predicts), it still has some loyal customers who will keep its returns reasonably high. On the other hand for a small firm, an absence of such customers can lead to sudden drop in revenue and therefore much greater drop in returns than for a big company. Hence we could see that asset's beta at various quantiles could be affected by other factors and this could be same in general for industries too. Their betas can be affected either by a size of the industry, dependency on exports or on political support. Realizing for which firms this occurs and in what scale can be important in investor's decision to buy or sell as this could be associated with massive earnings and losses. Our aim will be to investigate this problem and possibly identify these abnormalities.

However, we will start from the very beginning and in Chapter 2 we will introduce the CAPM. We will follow the approach of Sharpe, which seems to be more intuitive. In Chapter 3 we will discuss problems regarding the model, which assumptions are most often violated and why the model does not have empirical success. We will follow this discussion in the next chapter by analysis of real data. We will use Ordinary Least Squares (OLS) regression to estimate parameters and we will discuss if our data are in accordance with the CAPM. Chapter 5 will be devoted to the introduction of quantile regression, where we will describe what it estimates and we will show some asymptotic results and a test which is based on quantile regression. We will apply this methodology in Chapter 6 where we will investigate our main question, that is if asset's betas vary over return distribution.

Chapter 2

Establishing the CAPM

Building on the work of Dr. Harry Markowitz (Markowitz 1959), both William Sharpe (Sharpe 1964) and John Lintner (Lintner 1965) laid foundations of capital asset pricing theory by independently introducing the CAPM. These two economists used different approaches when deriving the model, with Sharpe presenting more straightforward and rather intuitive one. In the following pages, we will establish the aforementioned CAPM in the way William Sharpe did it in his paper.

2.1 The Minimum Variance Frontier

First, we will explain what the Minimum Variance Frontier (MVF) means, as it is important to understand this key feature in asset pricing theory. A good suggestion for a proper definition would be simply a set of all efficient portfolios, where efficiency is meant in a sense how everyone would intuitively imagine it. Roughly speaking, MVF is a boundary of a set of all feasible investment portfolios. However, to give it a proper definition, we should start with few important assumptions and specify what investment portfolio is.

By words investment portfolio is meant to be a combination of assets to which investors in the first period invest and from which in the second period get revenue. We assume that all investors view the outcome to have the same probabilistic distribution and are only interested in its expected value (expected revenue) and standard deviation (risk). Furthermore, we assume that all assets are risky and also infinitely divisible, which allows us to take into account not the absolute value of investment return but the rate. Moreover, we ex-

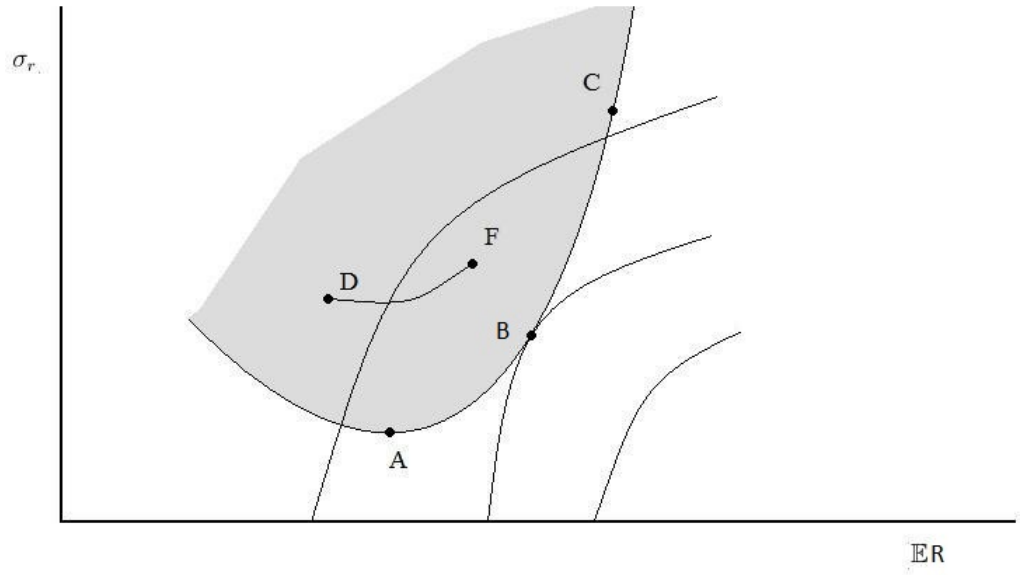


Figure 2.1: The Minimum Variance Frontier

Source: author's computations.

pect investor to act rationally, hereby prefer higher expected return rate to a lower value with constant variance and lower variance to higher with constant expected return rate. Let us denote $\mathbb{E}[R]$ the expected return rate and σ_R standard deviation of the return rate of an investment plan R. Let an investor has a utility function $U = f(\mathbb{E}[R], \sigma_R)$ which fulfils the above mentioned assumptions.

Now, consider $\mathbb{E}[R], \sigma_R-$plane and indifference curves of investor's utility function as shown in Figure 2.1. Every feasible investment plan is represented by a point on the plane according to its expected return rate and standard deviation. In the absence of risk–less asset, the set of available investment opportunities will be similar to the shaded area.

To get to know more about the nature of the set, consider two investment plans – D and F, and an investor who places a proportion of α, $\alpha \in \langle 0, 1 \rangle$, of his income in D and the remainder $1 - \alpha$ in F. Let denote H this new plan and ρ_{DF} correlation coefficient between rates of return of the plans D,F. Then, we know that

$$\mathbb{E}[R_H] = \alpha\mathbb{E}[R_D] + (1-\alpha)\mathbb{E}[R_F]$$

$$\sigma_{R_H} = \sqrt{\alpha^2\sigma^2_{R_D} + (1-\alpha)^2\sigma^2_{R_F} + 2\rho_{DF}\alpha(1-\alpha)\sigma_{R_D}\sigma_{R_F}}$$

The arc between points D and F roughly shows where investment plan H

may lie in the plane for all possible values of α in the case of correlation coefficient close to 1. If it is exactly 1, then we have $\mathbb{E}_{R_H} = \alpha\mathbb{E}_{R_D} + (1-\alpha)\mathbb{E}_{R_F}, \sigma_{R_H} = \alpha\sigma_{R_D} + (1-\alpha)\sigma_{R_F}$ so it follows that the arc would be a segment between these two points. On the other hand, when $\rho_{DF} = -1$, we would get for $\alpha = \dfrac{\sigma_{R_F}}{\sigma_{R_D} + \sigma_{R_F}}$ that $\sigma_{R_H} = 0$. Therefore there would exist a risk–less asset – a feasible point on the horizontal axis, which would violate our assumptions.

Now, as we understand the set of all available investment plans, we can define what MVF is. Firstly, a portfolio is said to be efficient, when its expected return rate is higher or equal than the expected return rate of every other portfolio with the same risk (measured by standard deviation). In our figure efficient portfolios are the ones on the right boundary of the shaded area. Minimum variance frontier is the set of all efficient portfolios, i.e. the curve ABC.

2.2 The Capital Market Line

Until now, we have been dealing with situations when no risk–less asset was available. Let us consider such an asset F, whose expected return rate $\mathbb{E}[R_f]$ is equal to the pure interest rate π. Moreover, we assume that there is unlimited borrowing and lending available at the interest rate π.

Consider an investor placing α of his income on the asset B and $1-\alpha$ on the risk–less asset F Note that in this case, $\alpha \in \langle 0, \infty)$, as values of α greater than 1 indicate that the investor has borrowed money. Using the fact that $\sigma_{R_f} = 0$. we can calculate the expected return rate and the standard deviation of this new investment plan G as follows:

$$\mathbb{E}[R_G] = (1-\alpha)\mathbb{E}[R_f] + \alpha\mathbb{E}[R_B],$$

$$\sigma_{R_G} = \sqrt{\alpha^2\sigma_{R_B}^2 + (1-\alpha)^2\sigma_{R_f}^2 + 2\rho_{BP}\alpha(1-\alpha)\sigma_{R_B}\sigma_{R_f}} = \alpha\sigma_{R_B}.$$

Clearly, for all values of α, plan G lies along the line FB In the case that we choose B such that the line FB is tangent to the MVF, as shown in Figure 2.2, we obtain a new set of efficient portfolios. The line FB, or more specifically the ray FB, is called the Capital Market Line (CML). It should be noted that it is required for the CML to be tangent to the MVF.

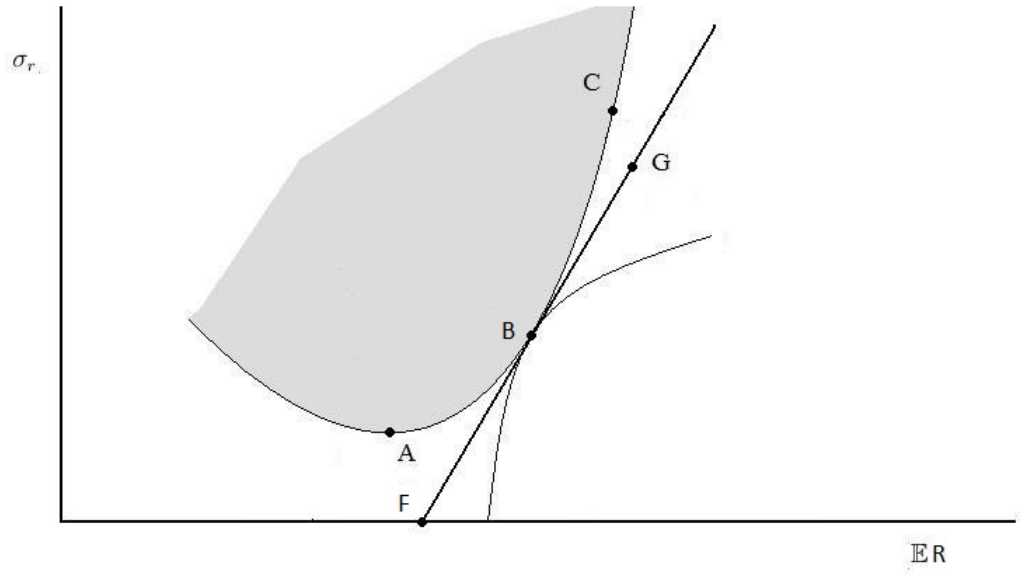


Figure 2.2: The Capital Market Line

Source: author's computations.

2.3 The Equilibrium

Recall the assumptions we set at the beginning of Section 2.1. We agreed that every investor faces the same distributional expectations. In other words, they must face the same MVF independently of their utility functions. Everyone can also use unlimited borrowing as well as the same risk free asset. To sum it up, desired portfolio of each investor lies on the CML and can be reached only by investing into the risk free asset and the unique tangency portfolio (In the case of Figure 2.2, portfolio B is the tangency portfolio) We call this unique tangency portfolio a market portfolio.

The core of the CAPM lies in the understanding what happens to the tangency portfolio when we change proportion of income invested in single asset. Consider the market portfolio M and a single asset A, which is part of M, as shown in Figure 2.3. Denote D a new portfolio with α of income invested in A and $1 - \alpha$ in M. Because some part of investor's wealth is already invested in M there exists $\alpha < 0$ such that it is still a feasible portfolio. For such α denote the portfolio C. Now, we have for every α from some reasonable interval an investment plan with the expected return rate and standard deviation as shown in Figure 2.3. Assume that the function of α, or curve AMC is smooth at the point M. Moreover, it is clear that the curve AMC cannot cross the CML

otherwise points on the curve, which are attainable, would be under the CML, therefore unattainable– contradiction.

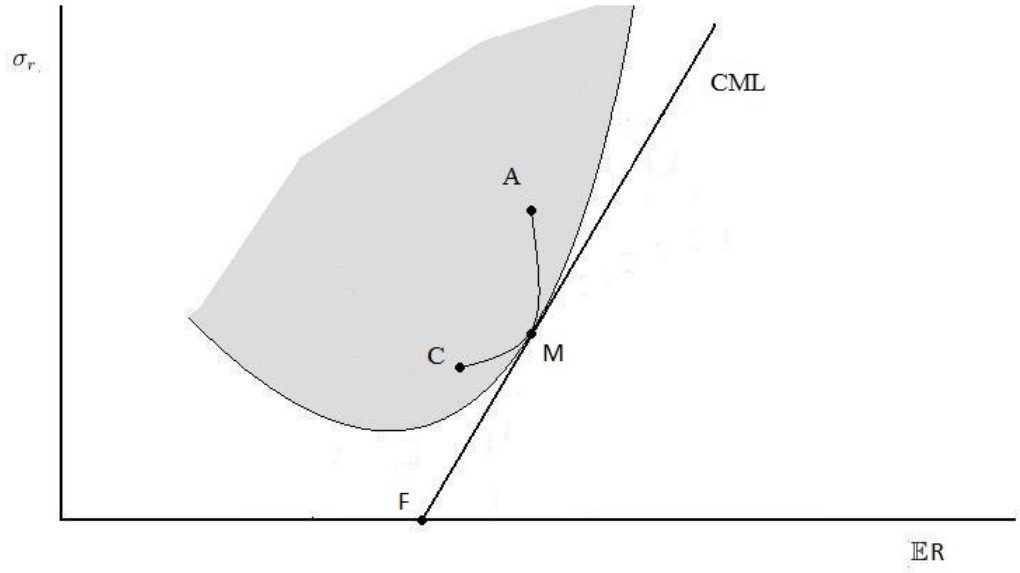


Figure 2.3: The Equilibrium

Source: author's computations.

Evidently, the slope of the curve AMC at point M must be equal to the slope of the CML. Using that

$$\sigma(\alpha) = \sqrt{\alpha^2\sigma_{R_A}^2 + (1-\alpha)^2\sigma_{R_M}^2 + 2\rho_{AM}\alpha(1-\alpha)\sigma_{R_A}\sigma_{R_M}},$$

we can derive that at point $\alpha = 0$, where $\sigma(0) = \sigma_{R_M}$,

$$\frac{d\sigma(0)}{d\alpha} = -\frac{1}{\sigma}\left(\sigma_{R_M}^2 - \rho_{AM}\sigma_{R_A}\sigma_{R_M}\right) = -\sigma_{R_M} + \rho_{AM}\sigma_{R_A}. \tag{2.1}$$

Similarly, $\mathbb{E}(\alpha) = \alpha\mathbb{E}[R_A] + (1-\alpha)\mathbb{E}[R_M]$, and for all values of α we have:

$$\frac{d\mathbb{E}(\alpha)}{d\alpha} = \mathbb{E}[R_A] - \mathbb{E}[R_M]. \tag{2.2}$$

Combining results from Equation 2.1 and Equation 2.2, we obtain:

$$\frac{d\sigma(0)}{d\mathbb{E}} = \frac{\dfrac{d\sigma(0)}{d\alpha}}{\dfrac{d\mathbb{E}(0)}{d\alpha}} = \frac{\sigma_{R_M} - \rho_{AM}\sigma_{R_A}}{\mathbb{E}[R_M] - \mathbb{E}[R_A]}. \tag{2.3}$$

The slope of the CML is $\frac{\sigma_{R_M}}{\mathbb{E}[R_M]-\pi}$, and as we know it has to be equal to the slope of the curve AMC at the point M, described by Equation 2.3. Solving for $\mathbb{E}[R_A]$ we obtain the result:

$$\frac{\sigma_{R_M}-\rho_{AM}\sigma_{R_A}}{\mathbb{E}[R_M]-\mathbb{E}[R_A]}=\frac{\sigma_{R_M}}{\mathbb{E}[R_M]-\pi},$$

$$\mathbb{E}[R_A]=\mathbb{E}[R_M]-\left(1-\frac{\rho_{AM}\sigma_{R_A}}{\sigma_{R_M}}\right)(\mathbb{E}[R_M]-\pi)=\pi+\frac{\rho_{AM}\sigma_{R_A}\sigma_{R_M}}{\sigma^2_{R_M}}(\mathbb{E}[R_M]-\pi).$$

Finally, we can use that the expression $\rho_{AM}\sigma_{R_A}\sigma_{R_M}$ is by definition covariance $\mathrm{cov}(R_AR_M)$. Plugging this into the last equation, it yields:

$$\mathbb{E}[R_A]=\pi+\frac{\mathrm{cov}(R_AR_M)}{\sigma^2_{R_M}}(\mathbb{E}[R_M]-\pi)=\pi+\beta_A(\mathbb{E}[R_M]-\pi), \tag{2.4}$$

where $\beta_A=\frac{\mathrm{cov}(R_AR_M)}{\sigma^2_{R_M}}$.

2.4 The Security Market Line

The last feature which will be discussed in this chapter will be the Security Market Line (SML). Let's recall the Equation 2.4 and solve it for β_A. We obtain

$$\beta_A=\frac{\pi}{\mathbb{E}_{R_M}-\pi}+\frac{1}{\mathbb{E}_{R_M}-\pi}\mathbb{E}_{R_A}.$$

The equation says, that for any single asset A which is part of the efficient investment portfolio M, its market beta depends linearly on its expected return rate. In other words, when we decide to invest in an asset with greater expected return rate, we can expect market beta of the asset to increase proportionally. That came certainly as a no surprise as market beta describes the magnitude of a risk which is correlated with the market and should be directly proportional to the expected return.

In the case there exists another tangency portfolio, it can be shown that the relationship with the market will be exactly the same. Moreover, we would find out that these tangency portfolios are perfectly correlated with each other and that market betas of individual assets would not change, as shown by Sharpe

(1964), pages 440-441. That allows us to call the tangency portfolio unique.

We have derived the traditional CAPM model the way William Sharpe did it in 1964. Concisely, the model says that the expected return rate of an asset A equals to the risk free rate π plus, reward for facing systematic risk – risk which is correlated with the market and can be predicted, of the asset, risk premium. Moreover, we learnt that a risk premium of an asset depends only on its covariance with the market. Finally the model also says that relationship between the expected return rate of an asset and its market beta is linear.

Chapter 3

Difficulties of the CAPM

In the previous chapter, we found out that the CAPM has very profound impact on behaviour of investors. As Fama & French (2004) noticed, it has three important implication stemming from the initial assumptions, which all can be subject of a test. Firstly, only asset's betas affect their expected returns with the relationship being linear. Secondly, the beta premium is positive, which means that the expected return on the market portfolio is greater than the expected return of assets uncorrelated with the market. Finally, the expected returns of assets whose returns are uncorrelated with the market returns are equal to the interest rate – risk free return rate. The reason is that beta for such an assets should be equal to zero.

3.1 Early Critics

Immediately after publishing the aforementioned model, the first tests were derived to confirm or reject the theoretical results. Both time-series regressions and cross-sectional regressions were used.

In cross-sectional regression, the approach was to regress average asset returns on estimates of assets betas. The theory then implies that the intercept is the risk free interest rate and the coefficient on beta is the difference between expected return of the market and the risk free rate. After overcoming two major problems with precision and bias of the data, the tests firmly rejected the CAPM. They found out that there is a positive relationship between beta and average return, however, it does not matched the predicted relationship.

The tests consistently found that the model underestimates the intercept and overestimates the coefficient on beta. This fits for number of early test, as noted again by Fama & French (2004), where some examples are presented on page 32, the second and the third paragraphs.

Time-series regression stems from the fact that the model implies the following relationship between expected return and market beta:

$$R_{it} - R_{ft} = \alpha_i + \beta_i(R_{Mt} - R_{ft}) + \varepsilon_{it}.$$

The fact that assets excess return is fully explained by its risk premium, which depends solely on beta and expected value of $R_{Mt} - R_{ft}$, means that the intercept term in the regression, called "Jensen's alpha," equals zero for every asset.

Time-series regression tests gave similar results as cross-section regression tests and confirmed that the relation between beta and average return is "too flat". For examples, see again Fama & French (2004).

Because of all these empirical failures of the model, its accuracy in estimation of expected returns of assets remains questionable. Economists argue, that this is because of its strict assumptions, which are not fulfilled in real situations. This leads to deriving new more complicated models such as Black CAPM, intertemporal CAPM and others.

3.2 Roll's Critique

Roll (1977) published a famous paper about testability of the CAPM. He concentrated on the observability of the market portfolio and made two key statements.

1. Mean variance efficiency is equivalent to the CAPM equation holding. That implies that for any given proxy of market portfolio, there is no difference between testing the CAPM equation or testing for mean variance efficiency of the portfolio, with both being equivalent.

2. The market portfolio has to include all available assets, including assets such as human capital, real estate etc. This means that the market

portfolio is unobservable and therefore returns on all possible investment opportunities are unobservable.

From these two statements, we can derive that the validity of the CAPM is equivalent to the market being mean variance efficient. However we cannot observe all investment opportunities, therefore we cannot test whether a portfolio is mean variance efficient. Hence because the tests use proxies and not the true market portfolio, we do not test the CAPM and we learn nothing about the CAPM.

This makes the CAPM empirically unusable because the concept of market portfolio lies in the heart of the model. Fama & French (2004) on page 41 state that the relationship described by the CAPM holds in any efficient portfolio, therefore it would be enough to find a market proxy that lies on the minimum variance frontier. However, the strong rejections of the CAPM indicate that no reasonable market proxy close to the minimum variance frontier was found. Moreover they add that if researchers are constrained to reasonable proxies, it is unlikely that they ever will.

3.3 Other problems

Apart from this rather fatal problem of the model, there are another assumptions, which can be called unrealistic. For example, model assumes that investors care only about mean and variance of one period return. There are clearly also investors who optimise their portfolio over long term horizon and others over short term horizon. Moreover we also assume that all investors have the same information and all know the true distribution of returns.

The earlier mentioned assumptions that investors care only about mean and variance of the return distribution is also extreme. It also makes sense that investors care about probability of extreme events not fully captured by variance or about other properties of the distribution, something which cannot be explained by two parameters. It is also rational that investors are concerned by how their portfolio co-varies with labour income, future investment opportunities or social status.

Another reason for empirical failures of the CAPM can be irrational pricing.

Behaviourists argue, that stocks with high P/B ration (share price of a company/book value per share) are expected to do poorly and on the other hand stocks with low P/B ration are expected to do well. Investors are influenced by past performances which causes higher prices for growth (low P/B) firms and vice versa. Few proponents of this view can be found in Fama & French (2004), on page 37 in the second paragraph.

It is well known that the CAPM has its flaws and that it has never been an empirical success. However, by deriving new capital assets pricing models its results can be dramatically improved. Fama & French (2004) proposed on pages 38-39 a new three factor model, which they argue performs much better than classical CAPM. Although added variables in the model are "brute force constructs meant to capture the patterns uncovered by previous work on how average stock returns vary with size and the P/B ratio" and have no theoretical explanation, it is not an obstacle in using this model in predicting expected return rates. That ultimately results in a model which captures more of the variation of expected returns of an assets.

As we could see, the CAPM does not appear to be working in practise and one could expect some empirical problems when fitting it. We will try to look at the problem of validity of the CAPM in a different way using quantile regression. The aim will be to detect differences in beta in different quantiles of return distribution, which could be caused for example by irrational pricing described in the previous paragraph. Bur first we will have a look at OLS regression where we will try to estimate the parameters of the CAPM in the classical way.

Chapter 4

Estimating the CAPM by Ordinary Least Squares Regression

In this chapter we will proceed to the application of the CAPM. First we will discuss what data we will use, and the reasoning for it, as it might not be that clear what is the return of market portfolio or other variables. Moreover, we will need to find a proxy for one parameter in our model – the risk free rate of return, as in reality there is no general agreement about a risk free rate, so we need to find a variable which approximates this rate. This will be all subject of discussion in Section 4.1.

In the second section of this chapter, Section 4.2, we will run Ordinary Least Squares regression in order to obtain estimates of market beta of estimated assets. We will also discuss whether these estimates can be considered as valid, as it will be needed to go through the all–important procedure of validating model assumptions. This, as we will see, will appear to be a limiting obstacle, mainly because of number of estimated models, which in fact could be much higher, which makes checking all model assumptions lengthy and complicated.

For estimating market β of various assets, we will use time–series technique which we mentioned in Section 3.1, where our model is in the form of

$$R_{it} - R_{ft} = \alpha_i + \beta_{iM}(R_{Mt} - R_{ft}) + \varepsilon_{it}. \tag{4.1}$$

To remind reader of the notation used, R_{it} stands for return of an asset i in time period t, R_{ft} denotes risk free rate in time period t, β_i is a market beta of asset i, α_i is "Jensen's alpha", which according to our theory should be zero

and finally ε_{it} is an error term, which we will assume is for given i independent and identically distributed across all time periods t. One more thing I would like to mention is the way how we can handle "Jensen's alpha", as if we would like to obtain estimates of market betas from the CAPM. First, we can see that if we run normal OLS regression with intercept, we obtain estimates for this "Jensen's alpha" and we can test whether it equals zero or not. If we reject this test for significance of this coefficient, we will have strong argument that something is wrong with our model. However, because of the theoretical derivation we would expect our data to comply with this theoretical result and then we can run regression without intercept to obtain "true"(according to our model) estimate of asset's market beta.

4.1 Description of the Data

In this section, we will introduce the dataset which we will use first for estimation of market beta of returns of portfolios which consist from firms in the same sector. These firms are assumed to have same market beta, as we derived in Section 2.3, beta of an asset i can be expressed as $\beta_i = \frac{\text{cov}(R_i, R_M)}{\sigma^2_{R_M}}$, where R_M is market return and $\sigma^2_{R_M}$ is its variance. Therefore given asset i market beta is determined only by covariance of return of this asset with return of the market, and it makes sense to assume that for firms in the same industry this covariance will be same. That implies, that market betas of firms in the same industry are equal.

Because we use time-series regression, we should also mention that over a short time period beta of an asset does not change. That allows us to consider market betas as a parameters which do not change, which is also one of the assumptions of OLS regression, and also of quantile regression, which we will run later. In our time–series, we will use five year monthly data starting 1$^{\text{st}}$ January 2009 and ending 31$^{\text{st}}$ December 2014. This choice was done as 6 years can be considered as relatively short time period, while monthly data will give us 72 observations, which should be enough to be able to make conclusions regarding our model. Moreover, a month between different measurements will be hopefully a gap big enough to ensure that our error terms are independent.

First, as I shortly mentioned at the beginning of this section, we will estimate market betas of portfolios consisting from firms of the same sector. For that we obviously need market returns of these portfolios, and for this purpose we will use data presented by K. French. This huge dataset contains observations from 48 industries in the last 90 years with number of different statistics and is available on-line at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. One of these statistics are portfolio's returns, which we use. We will also work only with data from the time period specified above.

We also discussed, that in our model we need an estimate (or possibly a true value) of risk free rate. Because the portfolios we will use consists from US companies, as a proxy for risk free return rate, I downloaded the return rate of 10 year US bonds which are possibly the safest option of investing money at US market, with relatively short maturity. Similarly for market return rate, we will use market index SP500 issued by Yahoo Finance which belongs between the most popular indexes of market performance of the US market and is calculated based on performance of 500 rather big companies across all sectors. From this market index one can derive the return rate by the usual increase divided by base value approach.

It is worth to note that data for our analysis were collected from different sources and were not prepared for educational purposes. Because of this, we might also need to be aware of some problems which arise with analysis of real data, we could for example mention influential points and/or points with high residuals which might make our analysis and resulting statistical inference invalid. Thankfully, in the data all values are stated so we do not need to think about what to do with missing values. Let us now go to the OLS regression.

4.2 Estimation of β

In this section, we will now be dealing with applying OLS regression to our dataset. As we mentioned before, we will try to justify the CAPM and to find out if our data follow this model. One important thing we will have to check for when making decisions about significant parameters in our model and its nature is if assumptions of the model are met. Especially in time–series regression, the fact that errors are correlated occurs relatively often and it almost

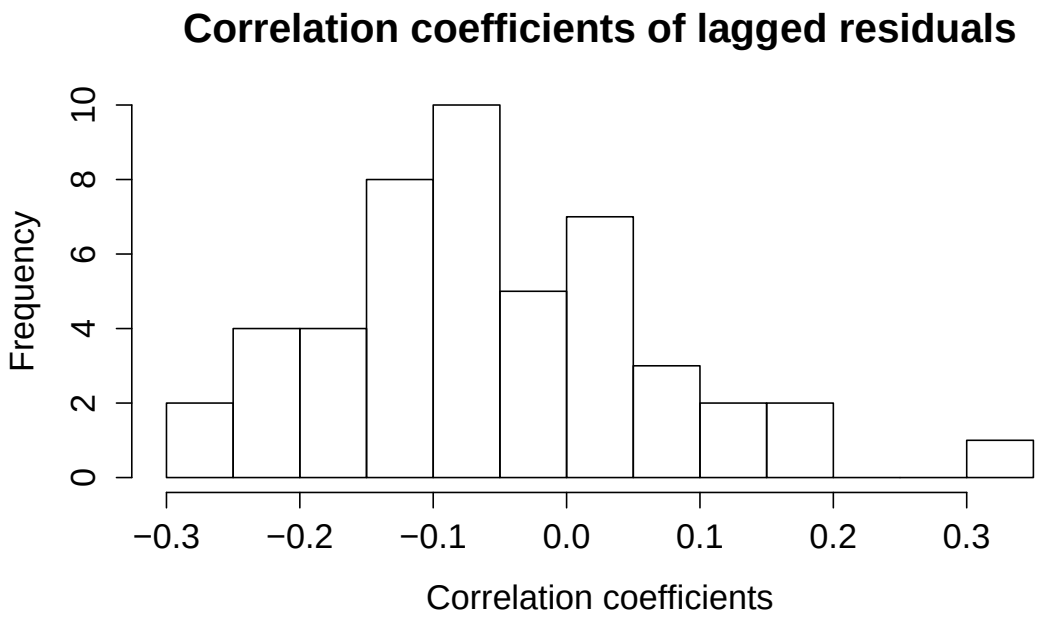
always has fatal consequences, as this correlation affects the statistical results in such a way that they do not hold, not even asymptotically.

First, we will examine model with intercept and we will test the hypothesis that the "Jensen's alpha" is zero for every portfolio. We should mention, that under this hypothesis, we will simply do t–test in OLS regression to determine if the coefficient is significant. However, because we will be doing this for 48 portfolios, even if our null hypothesis is true for all portfolios, we have $1 - (0.95)^{48} = 0.915$ chance, that we reject at least one of these hypothesis - i.e. we will make at least one type I error with probability of 0.915. Because that is rather a high number we should find a better way how to determine, whether our null hypothesis holds.

For that purpose, we will use another well-known fact and that is that if null hypothesis is true, then p-value of the test has uniform distribution on the interval $[0, 1]$. And because we will be doing 48 independent tests, if we plot a histogram of our p-values, we know that they should be somehow equally distributed around that interval. To give a proper mathematical conclusion to our test, we could run a Kolmogorov-Smirnov test for comparison of empirical distribution functions, which would tell us how close or how far we are from the expected distribution function of uniform distribution.

However, we should still pay attention to meeting model assumptions so we first examine correlation of lagged residuals to see if independence of errors can be assumed. Histogram of these coefficients can be seen in Figure 4.1 We see, that the correlation coefficients are not balanced and that there are more of them with negative sign, basically we could say that they are centred around -0.075 but with a bit heavier right tail. In absolute terms, the highest value of correlation coefficients of lagged residuals are errors from financial industry, which might represent some other problems in estimating beta for this industry. In general, other values of correlation coefficients are in norm as if we use Pearson test for correlation coefficient, then second lowest p-value for a hypothesis that there is no correlation between errors is 0.032 which with the number of tests we make is somehow in the expected region. I had a look also at Durbin-Watson statistics, whose values were somewhere in the expected region, so we can say that this assumption is more or less met in our data so we can run statistical tests and make statistical inference.

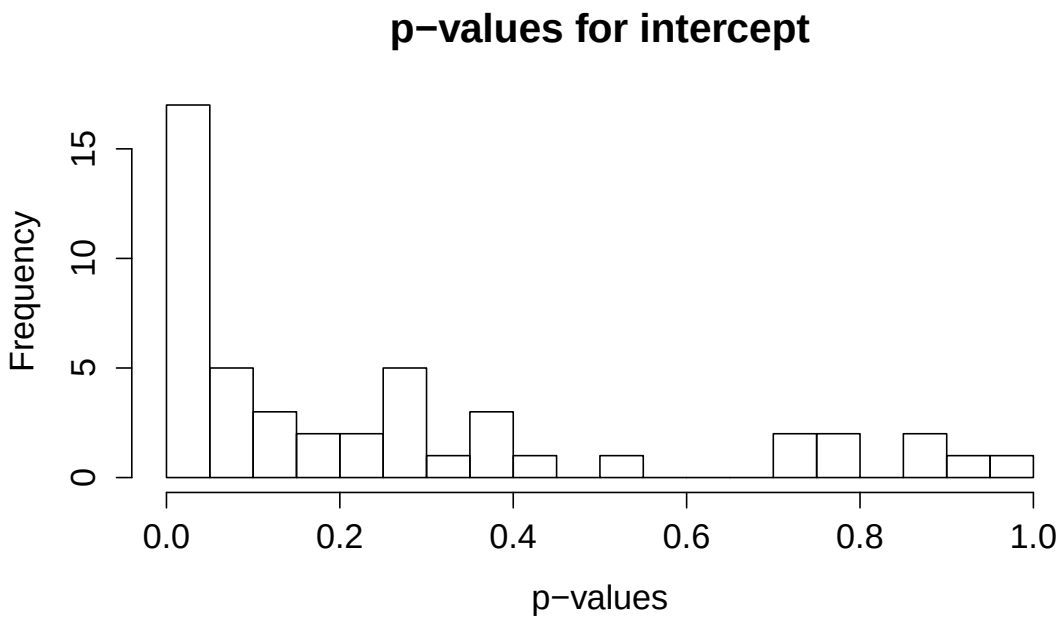
Figure 4.1: Histogram of correlation coefficients of lagged residuals in model with intercept.



With the methodology I mentioned at the beginning of this section, we run statistical test about hypothesis of significance for the intercept and present a histogram of p-values from these tests. That is presented in Figure 4.2. From there, we can see that distribution of these p-values is nowhere near to uniform distribution, which is also confirmed by the Kolmogorov-Smirnov test, which for comparison with uniform distribution gives p-value $3.035 \cdot 10^{-8}$, which suggest that we can reject our null hypothesis that intercept equals 0 for all industries. To stress how sure we are to reject this hypothesis, we have 17 p-values which are smaller than 0.05, with the smallest one having value of 0.0006. We can also explore in which way the intercept is fitted, if data suggest that if market return is equal to the risk free rate of return, then what kind of excess return we should expect from the portfolios, if lower or higher than risk free rate. That can be determined by looking at the t–statistics, whose sign and size say which way and by how much we are sure that intercept is smaller/greater than zero.

The answer on this question is clear, as only 10 t–statistics are smaller than 0, while 16 of them are higher than 2, which basically means that all

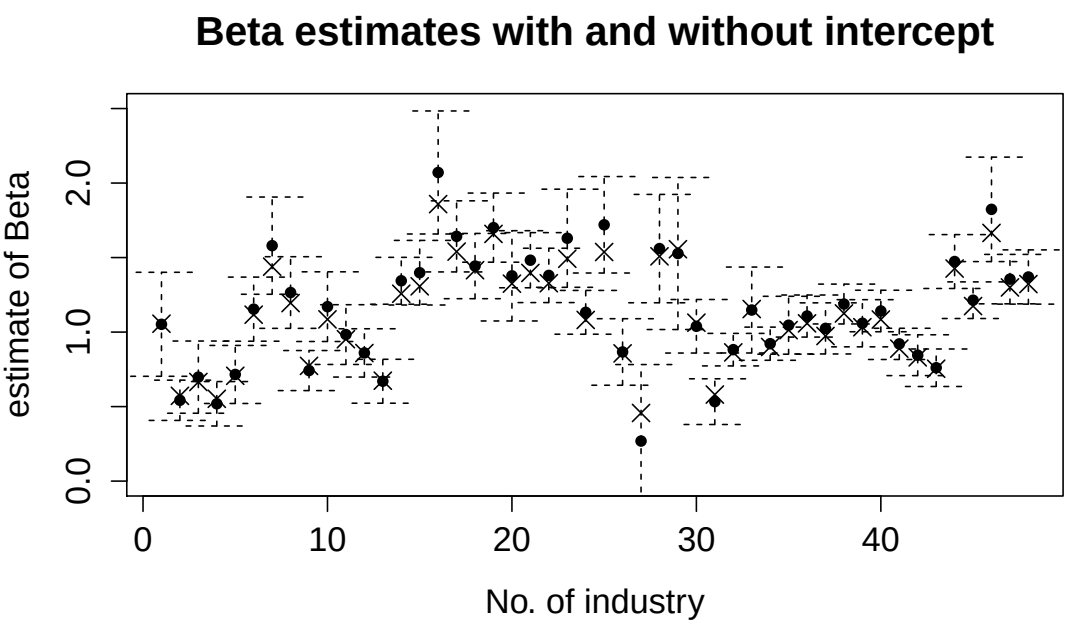
Figure 4.2: Histogram of p-values from test of statistical significance for intercept.



intercepts we would say are significant on 5% level are intercepts which are positive. We see, that this conclusions basically means that our theory of the CAPM is invalid, but I still feel that it might be useful to try to estimate betas in the model without intercept and to compare estimated values of betas in these two models together with their standard errors. In general, because we saw that intercepts we fitted were mostly positive, we would expect our estimates of betas to decrease (i.e. the line to get flatter), as most of our data have negative value of explanatory variable. Before we present the estimates of betas with and without intercept, we would like to add that we again checked for autocorrelation of residuals, which again suggested that there is no autocorrelation between them. Moreover, from the Durbin Watson test statistics, when we plotted histogram of p-values, we got histogram which was very close to histogram one would expect from uniform distribution, which is again in accordance with our null hypothesis.

That allows us to run OLS regression without intercept. In Figure 4.3, we present a plot of estimates of betas in regression with and without intercept, together with 95% confidence interval for beta in regression with intercept.

Figure 4.3: Estimates of betas(dots) and 95% confidence interval from OLS regression with intercept and estimates of betas(crosses) from OLS regression without intercept.



From that plot, we see that omitting intercept does not really affect the value of estimate of beta, which we are mainly interested in. All the estimates from regression without intercept are well in the confidence interval from the first regression. Similarly, if we looked at the confidence intervals from regression without intercept, we would find out that they are more or less same, apart from a small shift in the mean value.

It was rather interesting to see, that even though that used statistical tests on our model said that we should not omit intercept in our regression, we also found out, that adding the intercept in there does not give much of different estimate of the parameter we are interested in. This would somehow give more credibility to the CAPM, which says that the only thing which explains return of the assets are market betas together with the performance of a market. Therefore, the conclusion we can make from this analysis is rather unclear, as we basically stated that the CAPM is wrong, while we found out that it is not really much wrong. On the other hand, we did not check for other assumptions, both if errors are independently distributed and normal, so the conclusion of this analysis is not yet properly justified. Hence it might be useful to look at

other properties of the CAPM, which can be subject to test. One of them, as the title of this thesis suggests is that we might expect from various reasons that betas will vary in different quantiles of return distribution, which would be a contradiction of the CAPM. This is equivalent to errors being identically distributed. We should add that normality assumption is not that important as if without it results still hold asymptotically. We look at this question in the next two chapters.

Chapter 5

Introduction of Quantile Regression

In the following pages, we will be talking about regression technique called quantile regression, which is not that common as OLS regression and might be unfamiliar to the reader. We will state the model with parameters which quantile regression estimates. We will also mention some asymptotic results and present a test for homoscedasticity which uses quantile regression. This text will be based mainly on a book (Koenker 2005) by Roger Koenker, which presents a complete introduction to quantile regression and could be a good source to someone who wants to learn more about this interesting, powerful, and relatively unknown technique.

This method can be considered as a part of the big family of regression methods which estimate some parameter of a dependent variable given some other explanatory variables. The way it is constructed is a bit similar to classic linear models and the OLS regression which estimates mean of the return distribution. Quantile regression basically differs from OLS regression only by the parameter it estimates from the conditional distribution of the dependent random variable. Consequently, the process of deriving conditional quantiles is similar too and for easier understanding to the procedure of quantile regression estimation I recommend to realize what each step means or does in classic OLS regression.

5.1 Model Statement and Statistical Inference

Let us define a model where

$$Y = \beta_0 + \mathbf{X}^T\beta + U, \tag{5.1}$$

where $\mathbf{X}$ is a p-dimensional random vector, β vector of coefficients and U random error, such that without loss of generality with zero mean. This is a simple setting which we often meet in linear regression, but this time, we will not be interested in estimating the mean value of Y based on values of covariates $\mathbf{X}$ as usual, instead, we will focus our attention to quantiles of Y given values of $\mathbf{X}$. In other words, for $\tau \in (0, 1)$ and for some values of $\mathbf{X}$ we will be estimating conditional quantile $Q_y(\tau|\mathbf{x})$ of random variable Y.

For that purpose, we define a loss function as illustrated in Figure 5.1.

$$\rho_\tau(u) = u\Big(\tau - \mathbb{I}(u < 0)\Big).$$

Koenker (2005) in section 1.3 shows that if we have random variable Z and aim to minimize expected loss $\mathbb{E}\Big[\rho_\tau(Z - z)\Big]$, where z is a priori given number, then we choose z to be τ-quantile of Z. In other words if Z has got distribution

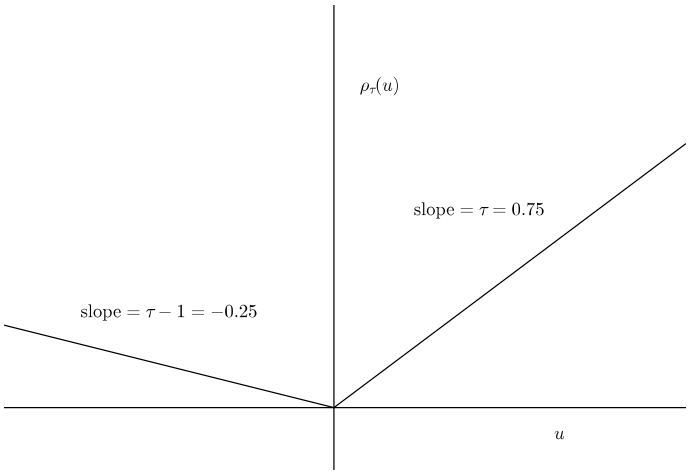


Figure 5.1: ρ_τ function shows weighting of absolute deviations, i.e. how they contribute to the final sum, shown for $\tau = 0.75$.

function G, then

$$\arg\min_{z\in\mathbb{R}} \mathbb{E}\Big[\rho_\tau(Z-z)\Big] = G^{-1}(\tau).$$

Given a random sample of $Z_1,\ldots,Z_n$ we can then minimize the sum $\sum_{i=1}^{n}\rho_\tau(Z_i - z)$, which will then give us an estimate of τ-quantile $G^{-1}(\tau)$. To extend this approach to quantile regression and the model stated in (5.1), assume that the random error term has distribution function F. Then we can express conditional quantile of Y given values of $\mathbf{X}$ as

$$Q_y(\tau|\mathbf{X}) = \beta_0 + \mathbf{X}^T\beta + F^{-1}(\tau), \tag{5.2}$$

hence from here we can see that the conditional quantile can be then modelled as a linear combinations of covariates and intercept. That implies, that we can model our conditional quantile of Y given $\mathbf{X}$ as

$$Q_y(\tau|\mathbf{X}) = \mathbf{X}^T\beta(\tau),$$

where β is a vector of our regression parameters, which we will want to estimate. Analogously to what we described in the case of single random variable, let us have a random sample $(\mathbf{X}_1, Y_1, \ldots, \mathbf{X}_n, Y_n)$, then we have U_i are *iid* and thus independent of $\mathbf{X}_i$. This assumption is needed as otherwise $F^{-1}(\tau)$ would not have been constant. Then estimation of coefficient $\beta(\tau)$ is done by minimizing the sum

$$\sum_{i=1}^{n}\rho_\tau(Y_i - \mathbf{X}_i^T\beta(\tau)),$$

which is a consequence of that conditional on $\mathbf{X}_i$ the expected loss $\mathbb{E}\Big[\rho_\tau(Y_i - \mathbf{X}_i^T\beta(\tau))\Big]$ is minimized when $\mathbf{X}^T\beta(\tau) = Q_y(\tau|\mathbf{X})$. Therefore, based on our random sample $Y_i, \mathbf{X}_i, i = 1,\ldots,n$ we can formulate our quantile regression estimate as

$$\hat{\beta}_n(\tau) = \arg\min_{\beta\in\mathbb{R}^{p+1}} \sum_{i=1}^{n}\rho_\tau(Y_i - \mathbf{X}_i^T\beta). \tag{5.3}$$

Although there does not exist closed form expression of this parameter estimate, finding the value of estimate can be done relatively simply by linear programming methods. We can see that the expression (5.3) is a convex function of β which can be rewritten to a linear function subject to some constraints. For these kind of problems we can for example use simplex method which is one of the most popular algorithms in linear programming. Details of the exact use

of simplex method to our quantile regression problem are described again in Koenker (2005), chapter 6. For our purposes it is enough to know that there exists efficient and quick method how to compute our quantile regression estimate.

However, estimation of parameters of models is only one part of the story in statistics. The other one, and possibly more important one is measuring how precise or how accurate our estimates are. In other words we want to find out how close to the real value we are with our estimate and maybe also what other values can be also considered as possible. Therefore we want to find distribution of our estimate in order to be able to construct confidence intervals and be able to make some conclusions on our hypotheses.

Assume that we have got random sample $Y_i, \mathbf{X}_i, i = 1, 2, \ldots$ such that conditional on $\mathbf{X}_i$ the distribution function of Y_i is F_i, let us also denote the conditional τ quantile of Y_i as $Q_{Y_i}(\tau|\mathbf{X_i}) = \xi_i(\tau)$. As stated and proved in Koenker (2005), page 120, to ensure some asymptotic properties of our estimator (5.3), we need to employ the following regularity conditions:

1. The distribution functions $\{F_i\}$ are absolutely continuous, with continuous densities $f_i(\xi)$ uniformly bounded away from 0 and ∞ at the points $\xi_i(\tau), i = 1, \ldots$.

2. There exists a positive definite matrix D_0 such that

   (i) $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T = D_0$,

   (ii) $\max_{i=1,\ldots,n} \frac{\|\mathbf{x}_i\|}{\sqrt{n}} \to 0$.

Denote then $\omega^2 = \tau(1-\tau)/f_i^2(\xi_i(\tau))$, where we can note that $f_i(\xi_i(\tau))$ is same for all i because of our *iid* error assumption and therefore ω^2 does not depend on i. Then under the conditions 1 and 2 we obtain

$$\sqrt{n}\Big(\hat{\beta}_n(\tau) - \beta(\tau)\Big) \to \mathcal{N}\Big(0, \omega^2 D_0^{-1}\Big).$$

Before we advance to discussing statistical inference which stems on this result, it might be useful to realize what these regularity conditions mean and what the D_0 matrix is. Basically in our *iid* error model setting the condition 1 means that our errors have absolutely continuous distribution and that at the quantile we estimate the density is finite and positive, which is rather a

weak assumption. The second condition places constraint on our explanatory variables, especially on the tails as we need a finite second moments of all covariates, as in case of existence $D_0 = \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]$ and rather light tails so condition (ii) holds too. However, in most applications these conditions looks to be met so the most important assumptions and much stronger assumptions which we stated was the one about *iid* error model and linear relationship, which is the one we will need to check for. We can also note, that under these conditions our estimate converges in probability to the true value of the parameter, which implies its consistency.

Because we now have a distribution of our parameter estimate, it is easy to construct confidence intervals for single parameter coefficients and also to construct critical regions for hypothesis testing. Or we can use Hotelling's T^2 test for testing for all coefficients in our vector of parameters. We will not discuss more this rather well known statistical theory and instead I will focus on presenting a rather new test for testing for heteroscedasticity, which is based on quantile regression.

5.2 Khmaladze Test for Heteroscedasticity

It is a common problem of linear regression that assumption about *iid* errors with constant variance is violated. However, assessing this assumption was, at least from my experience, always more heuristic than statistical, when after fitting the linear model one looked at residuals in order to check whether independence of errors and constant variance could be assumed. We can note, that in case of *iid* errors the quantile regression lines are going to be parallel as from the derivations we could see in (5.2) that the value of τ affects only the intercept. That gives us a hint how a test for heteroscedasticity could be constructed, as we can test whether all quantile regression lines are parallel to each other. Therefore we set our null hypothesis that all quantile lines are parallel and the alternative hypothesis is that this is not true.

As the derivation of the test statistics is rather cumbersome and uses advanced probability theory we will not be specifying it here. Its derivation can be found in Koenker & Xiao (2002a). The test is itself based on Kolmogorov-Smirnov convergence of empirical distribution function to the true distribution

function but with other present nuisance parameters it becomes rather complicated. The assumptions for this test therefore only require continuous one dimensional distribution of errors. To run the test in R we can use *KhmaladzeTest* function in R package Koenker (2015) named *quantreg*. To use it we specify the model relationship which we use. We also specify which type of test, whether "location" or "location-scale" hypothesis shall be tested. For our problem of heteroscedasticity it is the "location" version as we only test if the quantile regression lines differ by location. Finally, we need to specify set of τ values which will be used for testing the hypothesis. These need to be equally spaced. We shall also avoid trying to use too small or too high quantiles as these lines are often very imprecise as we do not have much information in our data about this region.

We make our conclusion on our hypothesis based on joint test statistic which is reported in the returned object under a header Tn. We compare this value the critical values calculated in Koenker (2005), page 318 or in the document Koenker & Xiao (2002b) available on-line at http://www.econ.uiuc.edu/ roger/ research/inference/khmal6ap.pdf. To find the right number we need to consider how much we truncated the true interval of quantiles $(0, 1)$, which means basically to set ε to the lowest quantile we used in the *KhmaladzeTest*. Finally, we set p as the number of slope coefficients in our model. Based on this asymptotic critical value we decide whether to reject null hypothesis or not. We reject null hypothesis if our reported test statistic is greater than the asymptotic critical value.

5.3 Applying and Interpreting Quantile Regression

One could see, quantile regression is rather a novel approach which can be used for modelling other aspects of dependent variable, not just its mean. Similarly to linear models, quantile regression also have non-linear and non-parametric versions. The one other application which stands out is when we have a model with heteroscedastic variance, such that variance depends on some linear combination of explanatory variables. Even in this set up quantiles are again just straight lines, who differ by shift and slope and in this case we can adopt quantile regression to estimate these values and again possibly make some statistical inference.

As I have already noted there exist an R package Koenker (2015) for fitting quantile regression called *quantreg* programmed by Roger Koenker. It offers few easily handled functions for fitting quantile regression and basic hypothesis testing for beginner users. The code uses simplex method for finding the quantile regression estimates and it works fairly quickly, as even in my computer when I fitted 48 times five quantile regression to dataset of 72 observations, it was done in one or two seconds. Apart from these basic functions which take care of fitting the model and following statistical inference the package also offers wide range of other functions associated with quantile regression which many advanced statisticians might find useful, these include fitting of censored quantile regression, bootstrapping quantile regression. I also find the help pages which are associated with functions I used very helpful and well written as the syntax follows the common-sense approach. Many references for deeper understanding are also given, even though most of the topics are well described in Koenker (2005).

Interpreting quantile regression lines might look like an easy task, however, there are few things we have to be wary of. First, because we often fit lines, they are expected to cross somewhere as even if they were originally parallel, the estimates will never be parallel (with probability 1). Koenker (2005), is aware of this problem and discusses it in Section 2.5, where we can learn that it is guaranteed that quantile lines do not cross at $\bar{x} = n^{-1} \sum x_i$. However, behaviour of the lines elsewhere is unpredictable and depends only on the data. Usually quantile lines then cross far away from the bulk of data, if the model is valid. In the other case we should pay special attention to our model, as there might be some serious flaws.

Finally, I would like to conclude this theoretical part with a short introduction of expectile regression, which in some sense can be seen as a similar and competing method for quantile regression and it might be good to be aware of it. Its difference to quantile regression is basically just that the argument in the loss function is a square of its value, which brings similarities to mean estimation, where we minimize sum of square errors. More on this topic can be found in section 2.8 of Koenker (2005). The main point we can learn in this part is however, that in contrary to quantile regression, expectile regression does not really have an easy interpretation as it is hard to say what exactly we estimate. Moreover, expectile regression also has one bad feature in the sense that if we

estimate upper expectile of the distribution, it is affected by how the lower tail of the distribution is distributed. One can feel that if we are interested in how distribution behaves in the upper tail, we should not really care what is in the lower tail, and that is what quantile regression does. From these two points of view, quantile regression looks more robust and more interpretable which makes it generally much more used method.

To sum up this theoretical chapter, we introduced quantile regression in mathematical detail in order to understand what it estimates and how it performs. Our main result which we will need in the analysis was then the Khmaladze test, which can be used for discussing heteroscedasticity of residuals. In this sense it could be considered as a competing method to assessing residual plots, as this is often problematic area of linear regression and it is area where quantile regression can help. Its advantage can be that it is carried out by computer code so we can do it for large number of models, in contrary to checking residual plot.

Chapter 6

Quantile Regression, Analysis of the Data

In this chapter, we will proceed with analysis of data presented in Section 4.1. Mainly, as the title of this thesis states, we will try to investigate whether there is any evidence of varying beta in different quantiles of the return distribution in our data set. We will explain how we can do this shortly.

First, let us review the statement of the CAPM, which can be written in the following form:

$$\mathbb{E}\Big[R_i\Big] = R_f + \beta_i\Big(R_M - R_f\Big), \tag{6.1}$$

where again we denote R_i return of an asset i, R_M overall market return and R_f risk free return rate and coefficient β_i, our market beta of an asset i. Similarly as in the case of the OLS regression, we will use time–series approach so we state a model

$$R_{it} - R_{ft} = \beta_i(R_{Mt} - R_{ft}) + \varepsilon_{it}, \tag{6.2}$$

where we assume errors ε_{it} are for given i independent for every t and errors for given i are also from the same distribution. Here, we can see that this model is somehow just a rewritten version of model in Equation (5.1) and therefore we can express conditional quantiles as in Equation (5.2) – i.e. they are just straight lines. Therefore, for given i if we use notation as was in previous chapter where $Y_t = R_{it} - R_{ft}$ and $X_t = R_{Mt} - R_{ft}$, then we have got random sample $Y_t, X_t, t = 1, \dots n$ from a distribution $R_i - R_f, R_M - R_f$ and we can express our conditional quantile model as

$$Q_{R_i - R_f}(\tau | R_M - R_f) = \beta_i(R_M - R_f) + F^{-1}(\tau), \tag{6.3}$$

where we defined $Y = R_i - R_f$ as a difference between asset return and risk free return, $X = R_M - R_f$ as an excess market return against risk free return and F is a distribution function of the error term. In this model β_i is still the asset's beta from Equation (6.1), therefore assumed to be same for all possible values of τ.

Now, let us consider what would happen if market beta of an asset i would vary in different quantiles of return distribution. Or in other words, for different values of quantile τ we would have different $\beta_i(\tau)$. That would result in a model

$$Q_{R_i - R_f}(\tau | R_M - R_f) = \beta_i(\tau)(R_M - R_f) + F^{-1}(\tau). \tag{6.4}$$

It is easy to see difference between models stated in Equations (6.3) and (6.4). In the first model, we have got for all values of τ the same slope coefficient β_i and different value of intercept $F^{-1}(\tau)$. Therefore, the quantile lines specified by this model are parallel. On the other hand, in model (6.4) we have for different τ different slope coefficient $\beta_i(\tau)$ and different intercept $F^{-1}(\tau)$, which implies that this lines are not parallel. This is therefore a simple consequence of possible varying beta over return distribution and as we saw it can be analysed by quantile regression. However, even if we have in our true model parallel quantile lines, the quantile regression estimates will not be parallel with probability 1. Therefore, to asses this null hypothesis if the lines are parallel we will need a statistical test, for that reason, we recall the Khmaladze test we discussed in Section 5.2, which deals exactly with this kind of problem.

As it was described earlier, for using Khmaladze test we need to choose a sequence of taus which will then be used for comparing the slope of the quantile regression lines, it is also worth to omit extreme quantile lines from our estimation as there is a big uncertainty resulting from only few data points on one side of the line. Therefore we use sequence of taus starting at 0.2 and ending at 0.8. Because it looks like that final value of the test statistics depends a lot on the choice of the sequence, we run test for three different choices $0.04, 0.05$ and 0.06. For this choice, we can learn that the critical values at $1\%, 5\%, 10\%$ level of significance, as calculated in Koenker (2005), page 318 or in Koenker & Xiao (2002b), are respectively $2.483, 1.986, 1.730$.

The results are presented in Table 6.1. We can note, that if we consider only mean value of these three test statistics, only in case of smoke (1.920), and bus

Table 6.1: Test statistics of Khmaladze test.

	0.04	0.05	0.06	min	mean	max
Agric	1.162	0.737	0.709	0.709	0.869	1.162
Food	1.032	1.203	0.648	0.648	0.961	1.203
Soda	0.827	0.624	0.687	0.624	0.713	0.827
Beer	1.199	1.166	0.682	0.682	1.016	1.199
Smoke	2.708	1.502	1.551	1.502	1.920	2.708
Toys	2.614	1.037	1.336	1.037	1.662	2.614
Fun	0.825	0.658	0.705	0.658	0.729	0.825
Books	1.055	0.870	0.674	0.674	0.866	1.055
Hshld	0.872	0.907	1.119	0.872	0.966	1.119
Clths	1.659	1.788	1.315	1.315	1.587	1.788
Hlth	0.468	1.680	0.760	0.468	0.969	1.680
MedEq	0.689	0.535	1.159	0.535	0.794	1.159
Drugs	1.129	1.927	0.944	0.944	1.333	1.927
Chems	0.753	0.626	0.765	0.626	0.715	0.765
Rubbr	0.725	0.934	0.503	0.503	0.721	0.934
Txtls	0.788	0.888	0.463	0.463	0.713	0.888
BldMt	0.854	2.085	1.001	0.854	1.313	2.085
Cnstr	0.748	0.614	0.848	0.614	0.737	0.848
Steel	0.622	0.686	0.928	0.622	0.745	0.928
FabPr	0.860	1.002	1.284	0.860	1.049	1.284
Mach	0.724	0.845	0.717	0.717	0.762	0.845
ElcEq	0.607	0.577	0.340	0.340	0.508	0.607
Autos	0.670	0.618	0.686	0.618	0.658	0.686
Aero	1.080	1.185	1.170	1.080	1.145	1.185
Ships	0.423	0.532	0.433	0.423	0.463	0.532
Guns	0.883	0.795	0.637	0.637	0.772	0.883
Gold	1.246	0.391	1.166	0.391	0.934	1.246
Mines	0.716	0.619	0.771	0.619	0.702	0.771
Coal	0.695	0.774	0.492	0.492	0.653	0.774
Oil	0.640	0.687	0.463	0.463	0.597	0.687
Util	1.700	1.399	1.177	1.177	1.425	1.700
Telcm	0.424	0.243	0.216	0.216	0.294	0.424
PerSv	0.656	0.754	0.332	0.332	0.581	0.754
BusSv	1.913	1.870	2.034	1.870	1.939	2.034
Comps	0.802	0.398	0.446	0.398	0.548	0.802
Chips	0.610	0.692	0.532	0.532	0.611	0.692
LabEq	0.571	0.400	0.723	0.400	0.565	0.723
Paper	0.617	0.507	0.823	0.507	0.649	0.823
Boxes	0.345	1.231	0.353	0.345	0.643	1.231
Trans	1.076	1.209	0.839	0.839	1.041	1.209
Whlsl	0.747	0.906	0.703	0.703	0.785	0.906
Rtail	1.026	0.963	1.024	0.963	1.004	1.026
Meals	0.522	0.571	0.632	0.522	0.575	0.632
Banks	0.934	1.168	0.886	0.886	0.996	1.168
Insur	0.600	0.398	0.781	0.398	0.593	0.781
RlEst	1.211	1.292	1.172	1.172	1.225	1.292
Fin	0.357	0.682	0.402	0.357	0.481	0.682
Other	1.429	1.562	1.342	1.342	1.444	1.562

service (1.939) we would reject null hypothesis at 10% significance level, but for neither one at 5% level of significance. However, what is worrying about these results is the great variability(some values differ by more than one) of the test statistics, which does not appear to have a good consequences, as if we could choose between any of the three Khmaladze tests we run on our data, we could consequently be able to reject null hypothesis for 4 industries at 5% level of significance and for 2 industries at 1% level of significance. the extreme example is toys industry, where in test with sequence 0.04 the test statistic was 2.614, while the test with sequence 0.05 scored 1.037, which is a value that does not allow us to reject null hypothesis at 10% significance level, while the former would lead us to reject null hypothesis even at 1%. It might be quite difficult to find an explanation for this kind of behaviour as these values were obtained from the same data. The reason why the test statistics for toys industry vary that much is not clearer after finding that sequences of slope estimates does not really differ, as presented in Figure 6.1.

Even though we do not consider standard errors of these estimates, one should expect for the same data with such a similar sequences of taus to get similar results and it is hard to find a reason why such a difference can occur. This appears to be a big flaw of Khmaladze test. To investigate the values of slopes of quantile regression lines of the industries whose hypothesis about parallel quantile lines we are more sure to reject, i.e. smoke industry and bus

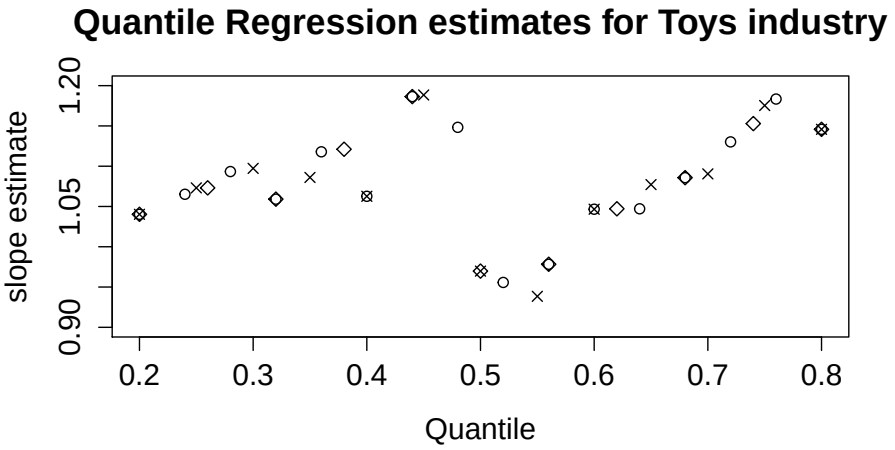


Figure 6.1: Quantile regression slope estimate for different quantiles. Circle denotes sequence of 0.04, cross of 0.05 and square of 0.06.

service industry we can see Figures 6.2 and 6.3. It is clear that in both industries the slopes of quantile regression lines show certain pattern which does not appear to be random which is reflected in the test statistics of the tests. Interestingly, the drop off looks to happen in similar quantiles, which might be, however, coincidence. To compare it with some other figures which do appear to have normal value of test statistic, we can see plot of slopes of quantile regression lines for telecommunication industry. This is showed in Figure 6.4

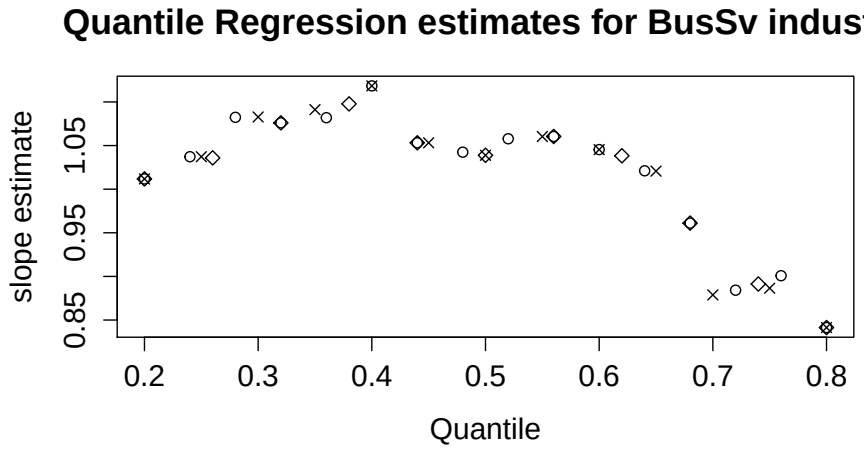


Figure 6.2: Quantile regression slope estimate for different quantiles. Circle denotes sequence of 0.04, cross of 0.05 and square of 0.06.

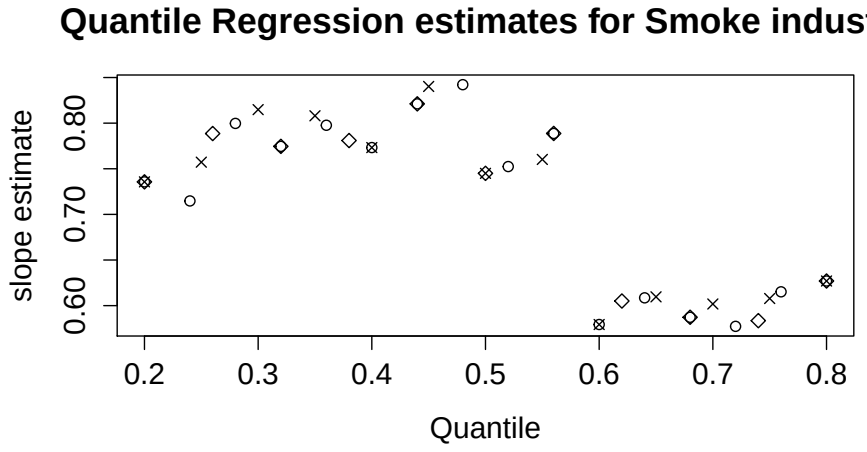


Figure 6.3: Quantile regression slope estimate for different quantiles. Circle denotes sequence of 0.04, cross of 0.05 and square of 0.06.

For the telecommunication industry, we can see that the slope values are distributed somehow randomly around value 0.87 and there is no reason to think that these values suggests that slopes of quantile lines are different. If we come back to the smoke and bus industry, one consequence of rejecting null hypothesis is that it means that the assumption about *iid* errors in the model is violated. It might be interesting to compare this result with the conclusion we would make if we were following the classic path of checking this assumption of the model in simple linear regression – fitted our model and then decided on its validity by assessing fitted values v residual plot. We present this plot for a smoke industry in Figure 6.5.

It would be rather interesting to see what would be the conclusion about the plot from a statistician as this plot is not too far away from what I would call ideal residual plot. I am sure that if I was given such a residual plot, I would say OK, this is fine. However, as we can see when we look at Figure 6.5 in more detail, there are more observations with positive residual which are close to 0 than on the negative side, which ultimately results in asymmetry which the test was able to capture. It is no coincidence that the other example where we would reject hypothesis at 10% level of significance for mean value of the three test statistics showed similar pattern of sudden drop of slope coefficients.

I was rather surprised that the test gave such a small value of test statis-

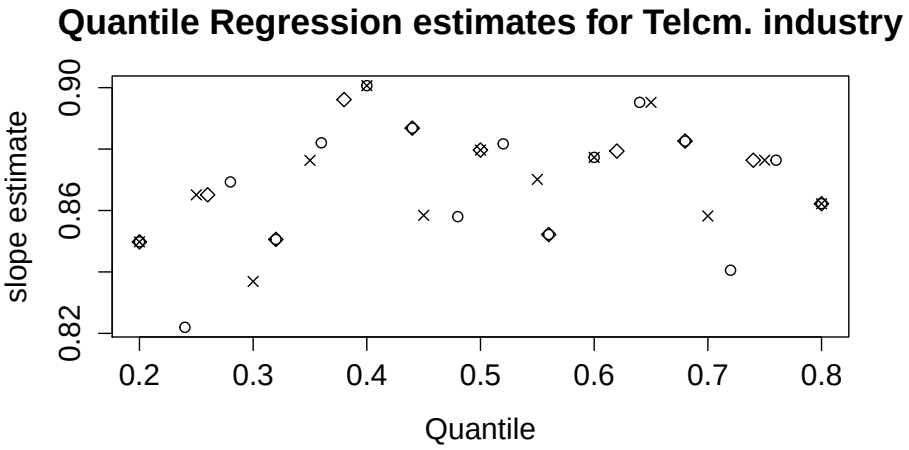
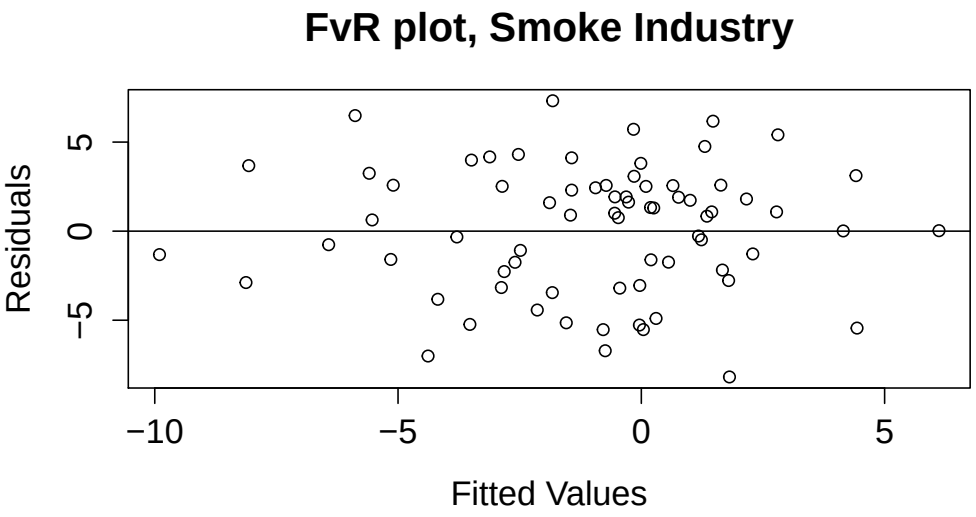


Figure 6.4: Quantile regression slope estimate for different quantiles. Circle denotes sequence of 0.04, cross of 0.05 and square of 0.06.

Figure 6.5: Fitted values and residuals plot from a simple linear regression on book industry.



tics for book industry, which appeared to have non-parallel quantile lines from the first picture I generated. Even in Figure 6.6 it is rather clear that the slope coefficients have increasing nature. The reason for this is that the slope estimate has too high standard error which results that for example the 95% confidence interval for median ($\tau = 0.5$) is circa $(1.06, 1.29)$ which means that

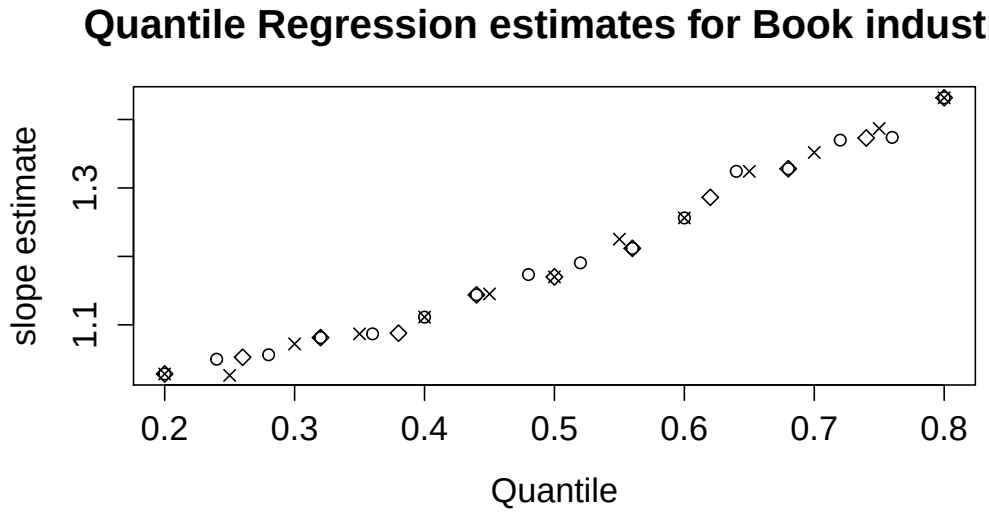


Figure 6.6: Quantile regression slope estimate for different quantiles. Circle denotes sequence of 0.04, cross of 0.05 and square of 0.06.

basically that there exists slope which is in all other confidence intervals for other quantiles. So although the quantiles show this increasing pattern, the uncertainty about our estimates is too big in our case so the result is that we cannot reject null hypothesis. Because of that, we should remember that slope estimates even with such a clear pattern do not tell us much about the validity of null hypothesis and even in this case the data collected might not suggest that null hypothesis should be rejected (although it "looks" obvious). That brings me back to our discussion about smoke and bus industry, where we argued about reasons why we rejected null hypothesis based on slope estimates. This discussion should not be interpreted in the way that if there is another plot like that, we reject null hypothesis. The main point I wanted to make was to give a reason why the test statistic was that high. On the other hand it is clear that this reason is not sufficient for test Khmaladze test to give that high test statistics.

Before I move to a part where we will discuss what we can do in order to correct for this problem of non–parallel quantile lines, I would like now to mention well–known problem of performing multiple statistical tests. When we make a decision about a statistical test on 5% significance level, it means if null hypothesis is true that with a probability of 5% we make type 1 error. If we are to do 20 independent tests, the probability of making at least one type 1 error is $1 - 0.95^{20} \doteq 0.64$ and in our case when we have 48 industries the probability of making type 1 error is 0.91, which is such a high number. If we were to control for this combined probability of making type 1 error for all test we would need to somehow lower our level of significance. One of the possible ways to do it is the Bonferroni method, which says that if we want to control combined probability of type 1 error to be α and if the number of statistical tests we do is m, then the approximate significance level we should use for rejecting single test is α/m. However, in our case, we would be having problems with calculating critical values, but because of the values of our test statistics and of the fact that not a single one does not exceed 2.8 while critical value for 1% level of significance is 2.483, it is very likely that after adopting the Bonfferoni method we would not reject any single hypothesis.

Even though that our data do not suggest that market beta varies in quantiles of the return distribution, it might be good to present a way how to correct the model so our analysis will be valid. By rejecting the null hypothesis about

parallel quantile lines, we basically state that variance of error terms is not constant. In that situation we need to find the way how variance changes and include this fact into the model. Here, we will present a way how to do it for variance which increases as a linear combination of explanatory variables. That would in other words mean the variance of returns changes linearly with excess market return. Using a general notation for explanatory variables $\mathbf{X}$ and dependent variable Y, then, we can define model:

$$Y = \beta_0 + \mathbf{X}^T\beta + \mathbf{X}^T\gamma \cdot U,$$

where γ is an unknown parameter and other parameters and variables are same as in (5.1). In this case, the conditional quantile can be expressed as

$$Q_y(\tau|\mathbf{X}) = \beta_0 + \mathbf{X}^T\beta + \mathbf{X}^T\gamma \cdot F^{-1}(\tau) = \mathbf{X}^T\beta(\tau),$$

which again means, that the conditional quantiles are linear. This is a very nice feature of quantile regression as the model for independent errors and model where errors depend linearly on $\mathbf{X}$ implies the same model for regression quantiles and therefore no matter which one of these models is valid we obtain the same estimate. Moreover, the Khmaladze test offers a version how to test whether the conditional quantiles are same up to location and scale shift, therefore if we apply this "location-scale shift" version of the test we can make conclusion if this model with errors which variance increase linearly with $\mathbf{X}$ is valid. The critical values are the same for this hypothesis too. With use of this test we can then make conclusions about both models and decide whether errors are constant, change linearly or change in a different way. Quantile regression in this case therefore presents a strong instrument which can help us to decide which model might be suited for our data.

Chapter 7

Conclusion

In this thesis, we introduced the CAPM and analysed a real data set by two regression techniques to obtain evidence if the CAPM holds or not. In the first part we run OLS regression to find out that one of the implication of the model does not hold, that is that the intercept in the time series regression appeared to be statistically significant. On the other hand, we were mainly interested in assets' betas, for which we found out that their estimates do not really change in both models as the imprecision of our estimates was much greater than difference in estimates of beta.

We also run quantile regression analysis in order to decide if our data suggest that the coefficient beta varies in different quantiles of return distribution. For this purpose, we used Khmaladze test which gave us an answer that this is not the case. We saw that economic interpretation suggested that for certain companies or even for certain industries, varying beta in different quantiles could have occurred. Because we worked with returns from portfolios of companies, we could not detect these firms and for example we could not decide if varying beta occurs only for firms with some common characteristics (size, past performance). However, we could decide about varying beta for some industries, but our data did not suggested that it is so. That was even in the case when the plot of betas suggested that there was clear increasing pattern. The case of smoke and bus service industry, which appeared to be the only industries where varying beta was possibly occurring, might have been only result of the number of industries we analysed. Obtained p–values from the tests we run were too small for us to be statistically sure that beta of that industry varies in different quantiles, as we wanted to control overall probability of Type 1 error.

We were in that case limited by our precision of estimates and maybe if we were able to obtain better estimates, our conclusion would have been different.

After all, we did not really used all the information provided in our data set, as we basically had panel data but we analysed it as a number of time–series regressions. Now, it is possible to analyse also panel data by quantile regression, applying this methodology on our data set, the story could have been different. Moreover, to increase precision, we could use longer time period, as this data are available to us. This represents a room for possible extension of quantile regression approach as this analysis would use all the information provided and would be therefore more effective.

Bibliography

Allen, D. E., A. K. Singh, & R. J. Powell (2009): "Asset Pricing, the Fama-French Factor Model and the Implications of Quantile Regression Analysis." *available at Research Gate* `http://www.researchgate.net/publication/49281596`.

Barnes, M. L. & A. W. Hughes (2002): "A Quantile Regression Analysis of the Cross Section of Stock Market Returns." *available at SSRN* `http://ssrn.com/abstract=458522`.

Chang, M. C., J.-C. Hung, & C.-C. Nieh (2011): "Reexamination of CAPM, An Application of Quantile Regression." *African Journal of Business Management* **5(33)**: pp. 12684–12690.

Fama, E. F. & K. R. French (2004): "The Capital Asset Pricing Model: Theory and Evidence." *Journal of Economic Perspectives* **18(3)**: pp. 25–46.

Koenker, R. (2005): *Quantile Regression*. Cambridge University Press.

Koenker, R. (2015): *quantreg: Quantile Regression*. R package version 5.19 - `http://CRAN.R-project.org/package=quantreg`.

Koenker, R. & G. J. Bassett (1978): "Regression Quantiles." *Econometric Society* **46(1)**: pp. 33–50.

Koenker, R. & Z. Xiao (2002a): "Inference on the Quantile Regression Process." *Econometrica* **81**: pp. 1583–1612.

Koenker, R. & Z. Xiao (2002b): "Inference on the Quantile Regression Process, Electronic Appendix." `http://www.econ.uiuc.edu/~roger/research/inference/khmal6ap.pdf`.

Lintner, J. (1965): "The Valuation of Risk Assets and the Selection of Risky Investment in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics* **47(1)**: pp. 13–37.

Markowitz, H. M. (1959): *Portfolio Selection: Efficient Diversification of Investment*. John Wiley and Sons, Inc., New York.

Roll, R. (1977): "A criticque of the asset pricing theory's tests' part i: On past and potential testability of the theory." *Journal of Financial Economics* **4(2)**: pp. 129–176.

Sharpe, W. F. (1964): "Capital Asset Prices: a Theory of Market Equilibrium Under Conditions of Risk." *Journal of Finance* **19(3)**: pp. 425–442.