LUND UNIVERSITY
School of Economics and Management

Default Risk in Equity Returns

A Study on Augmentation of the Three-Factor Model of Fama and French with Default Risk Factor

Tutor: Jens Forssbaeck

Authors: Aracelly Holst
Olena Martynenko

Abstract

Title: Default Risk in Equity Returns: a Study on Augmentation of the Three-Factor Model of Fama and French with Default Risk Factor

Seminar date: June 7th, 2010

Course: Master Degree Project in Finance, 15 ECTS, BUSM65

Authors: Holst, Aracelly and Martynenko, Olena

Supervisor: Forssbaeck, Jens

Key words: *ASSET PRICING*, *DEFAULT RISK*, *FACTOR MIMICKING PORTFOLIO*, *FACTOR RISK PREMIUM*

Purpose: The current thesis assignment aims to quantitatively verify systematic character of default risk and the statistical quality of the competing three- and four-factor asset pricing models.

Method: The experimental design applied to this study is premised on the three-factor model of Fama and French enhanced by default risk factor. The study utilizes the factor mimicking portfolio technique for modeling the risks underlying size, value and default risk factors. Distance-to-default estimate, deduced from the option-based model, is adopted by this study as a proxy for default risk. Regression analysis is applied on the time series of average returns on the portfolios of stocks possessing the pre-specified corporate characteristics.

Conclusions: The augmentation of the three-factor model with default risk factor improves the performance of a conventional asset pricing specification on average. The factor loadings of the portfolios of size, value and default risk factors exhibit strong properties of risk factor sensitivities for stocks. However, this holds only for the model that explains the average returns on the portfolio of stocks characterized by high value of market capitalization, high value of book-to-market equity, and short distance-to-default estimate. The size and value factors are found to be common in equity returns, but at the same time not being proxies for default related information. The study provides no evidence for the default risk being proxy for sensitivity to common risk factor in returns.

Table of Contents

1 Introduction

The current study aims to align two research areas of financial economics – asset pricing modeling and default risk estimation. Whether default risk is a systematic risk factor and therefore is common in equity returns, is an interesting question for the investing public. This chapter contains a background to the problem of default risk in equity returns and positioning of the current study, which allows for a clear formulation of research questions. A purpose of the study, delimitations and weaknesses are also presented in the chapter.

1.1 Background

The stock return development over time may be sensitive to a variety of factors, such as macroeconomic variables, idiosyncratic information, and politics, if speaking generally. Standard asset pricing models, such as CAPM of Sharpe (1963, 1964) and Treynor (1961), Arbitrage Pricing Theory (APT) of Ross (1976), and Three-Factor model of Fama and French (1992, 1993), emphasize that only systematic risk factors affect stock returns.

The conventional asset pricing models exploit the “ideal world” assumptions such as perfectly diversified investors with homogenous market expectations (CAPM) or arbitrage conditions (APT). Hereby, these models account for the risks that are common in stock returns and hence non-diversifiable. Systematic risks have empirically been confirmed to be commanding significant risk premiums, implying that they are priced in the market.

However, stock price/return development over time reflects a wide spectrum of market information, not least market reactions on changes in unique characteristics of the corresponding firms. Firm specific information is actually a source of idiosyncratic risks investor is exposed to. These risks are uncorrelated with systematic ones and may be diversified away through portfolio holding.

Thus, investor expectations are subject to both macro environment (including industry conditions, international economic conditions, and politics) and performance of a particular company (which is, in turn, dependent on a macro climate). Uncertainty and/or unanticipated changes in industry and company performance form the investor expectations about future patterns of stock price and return. Investor expectations, formed under ascending uncertainty, may negatively affect stock liquidity with implied limitations and inefficiencies for investors, leading to a higher required rate of return and thus determining a future development of stock price and return. It is therefore of importance to determine the potential risk factors, and hence sources of uncertainty, and their possible influence on investor expectations and market development.

While the conventional asset pricing framework states that abnormal returns of risky assets are attributable to bearing market risk solely, there is considerable evidence against the empirical and theoretical robustness of the classical asset pricing models. A number of studies suggest a variety of additional variables that command significant risk premiums. An Intertemporal CAPM theory of Merton (1973) is one of the earliest examples that can be mentioned in connection to this issue.

Risk factors that command significant risk premiums are the factors that systematically affect the returns of risky assets, and hence are those that represent non-diversifiable risks, investor would like to hedge against. (Chen *et al*, 1986) Empirical results of the study of Chen *et al* (1986) show, in favor of the CAPM-theory, that the market portfolio really explains a large portion of stock return variation. At the same time, this study provides evidence of insignificancy of market portfolio when simultaneously controlling for other risk factors. The three-factor model of Fama *et al* (1992, 1993) is another well-known example on this field.

Therefore, it is both of interest and importance for the investment world to determine, which economic risk variables, beyond the market risk factor, do command significant risk premiums and hence are systematic risk factors. The proper selection of the relevant risk factors is highly important. It is of a particular interest to determine the risks that are originated both by market uncertainty and idiosyncratic information, and which are able to explain the variation in stock characteristics over time.

Assessment of the portion of systematic riskiness in the risk factors closely connected to business uncertainty (idiosyncratic information) may lead further to development of qualitatively new area in asset pricing modeling. It may be true that the conventional risk factors are actually proxies for some more sensible, business-connected circumstances. Determination of these proxies would mean elaboration of asset pricing specifications that relax the strong assumptions underlying the conventional asset pricing models and/or provide a sensible explanation for the Fama-French factors. A possible example of the risk factor that is closely business-related and at the same time may contain a portion of systematic riskiness is a default risk factor.

Investigation of some proxies for default risk has documented that default risk contains idiosyncratic information and, disregarding this fact, also is highly exposed to changes in macroeconomy and industry. (Bonfim, 2009; Pesaran *et al*, 2003; Qu, 2008) The studies on extension of the conventional asset pricing models by different default risk variables show no agreement what regards the obtained empirical results. That is so despite the fact that modeling specificity of the default risk variable is often very similar from study to study.

1.2 Positioning of the Current Study

The classical asset pricing specifications have been a subject for criticism not entirely because of their equilibrium assumptions, but also because of their choices of explanatory factors. For instance, CAPM (Sharpe, 1963) unrealistically contemplates a single market factor as a proxy for all systematic risks; the Arbitrage Pricing Theory of Ross (1976) extends the CAPM by additional risk factors beyond market risk factor, but still considers solely the states of economy (systematic risks) assuming that investors are fully diversified; and the Three-Factor model of Fama *et al* (1992,1993) rests on empirically chosen variables, which have been shown to be proxies for other, more explicable, factors.

The traditional asset pricing specification of the three-factor model can be found in the contemporary literature being a subject for revising. To do so, the researchers choose, among other variables, default risk as additional model component.

In general, the existing research on this issue can be grouped in accordance with default risk proxy employed. Thus, the two main classes of studies can be detected: studies that use a default risk measure retrieved from the structural models, such as option-based models

(Gharghori *et al*, 2009); and studies that use systematic fraction of default risk from the marketable measures such as corporate bond spreads (Anginer *et al*, 2010) or spreads of debt obligations on credit derivative indices (Chan-Lau, 2006).

The studies on the field of interest differ not only by their choices of proxies for default risk factor. There also exist some differences regarding working hypotheses elaborated. There can be defined a group of studies aiming to investigate how the classical asset pricing specifications perform after their augmentation with default risk factor (*e.g.* Vassalou *et al*, 2004); whereas another group of studies addresses an issue of relationship between default risk and equity returns focusing on defining default factor risk premium (*e.g.* Ferson *et al*, 1991). Of course, one may say that both groups consider an optimal model design in the first place. Nevertheless, each of the groups of studies communicates somewhat different objectives and hence uses different statistical tests – either testing the overall modeling quality or specific interplay between variables in question.

The current thesis is generally based on a prior research. It concerns the asset pricing set of problems in relation to information on defaulting. A starting point of this study is a traditional specification of asset pricing, based on the three-factor model of Fama *et al* (1993). Mimicking portfolio formation procedure follows this classical study as close as possible but with respect to resources and tools availability.

Further, the model is augmented with default risk factor in order to control for appropriateness of model specification. Simultaneously, it is controlled for whether default risk is priced in the market and common in equity returns through defining the sign and significance of the default factor risk premium. Distance-to-default is employed as a proxy for default risk and is extracted from the option-based model. So far, the current study replicates to a relatively large extent the existing literature on this issue.

The novelty of the current thesis is in the methodological approach for the default risk factor construction. Variation in estimates of distances-to-default is used for building two mimicking portfolios (one mimicking portfolio for each of the competing three- and four-factor asset pricing models) that stand for the background default risk factor. Some prior studies also use distance-to-default estimate as a proxy for default risk, but in these cases a variable of aggregated default probability is generated from the estimated distances-to-default and used in the regression analysis. (Vassalou et al, 2004; Gharghori *et al*, 2007) Unlike these studies, the current thesis chooses to mimic the risk underlying default risk factor.

The mimicking technique provides the corresponding mimicking portfolio with ability to capture the information in default risk factor that is predominant for the examined asset returns, which reduces the noise. Another reason for preferring this method is in mitigation of the problem of errors in variables and no requirements what regards variables normalizing. (Huberman, 1987) Implementation of this technique allows determining whether mimicking portfolios are able to capture the common risk factors in equity returns. If so, the factors are systematic risk factors and are priced in the market.

1.3 Problem Discussion

As investor expectations are subject to both macroeconomy and firm related information, the distance-to-default measure is an appropriate proxy for default risk factor when explaining equity returns within the expanded three-factor model. Distance-to-default provided by structural models (Merton, 1974, Vasicek, 1984, Brockman *et al*, 2003) is a measure of default risk that incorporates important corporate information such as capital structure and asset volatility, and which is reactive to economic environment.

Besides, this measure of default risk concerns a wide spectrum of corporate decisions – from investment and capital structure decisions to possibilities of re-negotiation upon distress. This is so because changes in capital structure provoke changes in interplay between essential elements of the structural relationship that generates distance-to-default estimate.

Thus, leverage level matters, and its optimal level implies, among other things, keeping a safe distance from the default point. An upward departure from the optimal capital structure implies actual increase of WACC and thus – through some intermediary factors – increase of default risk. The more expensive capital is a consequence of credit spreads widening and hence worsening of credit quality, reduction of tax shield value, profits fall and interest expense rise. (Ogden *et al*, 2003; Pettit, 2007) Although it is not very clear what comes first – deterioration of operational performance or financial constraints connected to investment opportunities and returns on investment (Gertner *et al*, 1991; Wruck, 1990), there must exist a relationship between leverage and distance-to-default estimate. However, this relationship seems to be quite opaque and to a great extend depending on how asset market value is influenced by changes in capital structure.

Also, an overall business uncertainty can influence distance-to-default estimate primarily through asset volatility, which is a substantial component of structural default risk modeling. Higher asset volatility implies higher volatility of firm value and hence higher volatility of stock. Then, the investor uncertainty about trading asset's value makes her to take a relatively passive trading position which would induce the stock price to decrease.

Even a firm's potential ability to exploit financial restructuring tools and re-negotiate with creditors under a threat of distress/default may add to understanding of a rather ambiguous relationship "default risk – expected stock return" and historical development of distance-to-default estimate. Although it is reasonable and theoretically sound that higher risk implies higher return, it must not always be true for a distressed stock and probably depends on a level of financial distress the company is living through.

Additionally, it is reasonable to expect default risk to be common in stock returns due to the portion of systematic riskiness in the default risk factor. Indeed, there are studies that document significantly positive default risk premiums (Chen *et al*, 1986; Ferson *et al*, 1991) indicating systematic character of default risk. These studies, nevertheless, utilize corporate bond spread as a proxy for default risk, which is argued to be closely related to the market. (Demchuk *et al*, 2005) On the contrary, however, there exists evidence of corporate bond spreads not being originating positive risk premiums and thus not explicitly accounting for systematic component of default risk. (Anginer *et al*, 2010)

When employing distance-to-default as a proxy for default risk, a non-linear relationship between stock return and default risk is quite sensible. This default risk proxy is closely related to the firm specific characteristics and possibility of recovering from distress. Then, deeply distressed equity is rather unlikely to generate abnormal returns in response to information about "incurable" distress and hence extremely negative investor expectations. At the same time, the return reaction would be different when a recovery due to, say, changes in financial structure is feasible. Summarizing, it can be stated that higher default risk earns higher return; however, it must not hold as a firm approaches default point.

In addition to establishing the relationship between default risk and average equity returns within the asset pricing modeling, verification of whether default risk is a systematic risk factor and therefore is common in equity returns is within the scope of this study. In other words, the current study aims to test whether default risk originates a positive and statistically different from zero risk premium, which is basically a value of insuring against non-diversifiable portion of default risk. Absence of consensus what regards empirical findings on the "default risk – stock return" relationship and also on whether default risk is commonly priced in stocks, makes the investigation of the three-factor asset pricing model enhanced by default risk factor interesting in particular.

1.4 Purpose and Research Questions

With respect to the afore-presented discussion, the purpose of this thesis assignment can be stated as follows:

to verify quantitatively the systematic character of default risk and the statistical quality of the competing three- and four-factor asset pricing models.

The following research questions serve to refine the stated goal:

1) Does the augmentation of the three-factor model with default risk factor improve the performance of the model? Have the size and value factors been found concentrated in default risk and hence lose their explanatory power as regards equity returns when the default risk factor is included in the model?

2) Does default risk factor exhibit an explanatory power as regards equity returns? Does default risk possess statistical properties of a systematic risk factor and hence is priced in the market?

1.5 Delimitations and Weaknesses of the Study

The current study analyzes the explanatory power of the four-factor asset pricing specification in relation to the three-factor model, using the stock returns of the randomly selected companies listed on the NASDAQ stock exchange. The object of investigation is delimited to the NASDAQ non-financial companies due to incomparability of capital structures of financial and non-financial businesses. As such, the quantity of the working sample equals 171.

The explanatory variables used for analyzing the variation in average equity returns over time are delimited to the three factors adopted by Fama and French in their three-factor model (market risk factor, size and book-to-market equity) and the default risk factor.

The time horizon for this study is set between February 1991 and January 2010. The sample period is primarily based on the availability of data. The regression estimations are conducted on a monthly basis. However, due to financial reporting frequency, the distance-to-default estimates are calculated on a yearly basis.

The possible weaknesses of this study are judged to be originated by the sample size. The descriptive statistics for the time series of portfolio returns could be more desirable, and the regression inferences could be therefore more reliable, if the working sample also contained the firms listed on other stock exchanges in the U.S. in addition to the NASDAQ. The chosen sample size is restricted to 171 firms due to time constraints, since the sorting process, portfolio formation, and working up the input data matherial for the distance-to-default estimation are rather time consuming.

However, the authors are aware of that the larger sample would imply larger and more portfolios, and hence better mirroring of the real life variety in returns' characteristics. This, in turn, would imply lower correlations between the time series of returns on factor mimicking portfolios and consequently more reliable regression inferences. The number of dependent portfolios used in the study exceeds the number of explanatory mimicking portfolios. As such, a variation in firms' characteristics can be considered to be captured by the dependent portfolios. Unfortunately, the sample size used in this study cannot provide an opportunity for constructing even more dependent portfolios.

Another possible weakness of this thesis can be attributable to the distance-to-default variable. The problem is once again connected to the sample size. Since the distance-to-default estimation method is built on assumption of normally distributed asset returns (which is academically referred), instead of using empirically obtained default frequencies, a larger sample would probably provide more desirable descriptive statistics for the time series of the averages of the distance-to-default estimates.[1]

In addition, this study exploits a deductive approach, which implies that the research process exists within a certain theoretical framework that influences both modelling and conclusions. Moreover, the variables employed in the regression analysis are specified theoretically, so there is a possibility of theoretical predetermination of the findings.

1.6 Thesis Outline

Chapter 2 provides a theoretical background to the thesis problem. First, the three-factor model of Fama and French and empirical findings of existing studies on default risk in equity returns are presented. Thereafter, the structural credit risk modeling is discussed. *Chapter 3* explains the chosen approach for distance to default calculation, dependent portfolios formation, mimicking portfolios modeling, and factor risk premiums estimation and testing. That is done alongside with generation of the working hypotheses regarding relationship “distance to default – average stock returns” and whether the size and value factors are proxies for default related information. *Chapter 4* contains descriptive statistics of the time series employed and the model estimation results. *Chapter 5* analyzes the empirical results of the study with respect to theoretical framework and working hypotheses. *Chapter 6* contains conclusions and some further research directions on the field of study.

[1] The descriptive statistics of the time series of distance to default can be found in *Appendix I, Table I.1*

2 Theory

In this chapter the theoretical background to the thesis problem is presented. First, the three-factor model of Fama and French and the findings of existing studies on the area of interest are introduced. Thereafter, the chapter discusses structural credit risk modeling. The chapter aims to clarify for a potential reader the theoretical grounds that are essential for understanding of the thesis approach.

2.1 The Three-Factor Model of Fama and French (1993)

The paper by Fama *et al* (1993) aims to identify five common risk factors – the overall market risk factor, size (market equity, which equals to stock price times number of shares), value (book-to-market equity), leverage and earnings-to-price ratio. These variables are determined empirically and seem to explain the average returns on stocks and bonds. Exploiting the idea of strong integration between stock and bond markets, the authors examine whether the variables that explain bond returns are also important for stock returns prediction, and *vice versa*.

The time series regression results indicate that the factor loadings on the size and value factors, modelled as factor mimicking portfolios, exhibit strong properties of risk factor sensitivities for bonds as well as for stocks. Moreover, the time series regression estimations provide evidence on the following asset pricing issue: the size and value variables act as proxies for sensitivity to common risk factors in returns; thus, the risks underlying size and value variables are to be considered as systematic.

In relation to the current thesis, it is of interest to illustrate the empirical modelling applied by Fama *et al* (1993). Since this thesis investigates the time series of the average stock returns, the explanatory variables that are directed at stocks (namely, the size and book-to-market variables) are in the focus of this subsection.

Fama *et al* (1993) mention, that size and book-to-market are related to economic fundamentals. Thus, the firms with high book-to-market ratios (implying low stock prices relative to book values) are shown to have persistently low earnings on assets at least five years before and five years after the time point the ratio is calculated at. Regarding size, small firms are shown to have lower earnings on assets compared to big firms. Then, size might be associated with some common risk factor that explains the negative relationship “size – stock return”. What regards the value-variable, it is associated with relative profitability as a source of common risk factor in returns, which may explain the positive relationship “value – stock return”.

It is worth noticing, that when constructing mimicking portfolios, Fama *et al* (1993) pursue the following reasoning. Since the value-variable is expected to have a stronger role in average stock returns in comparison to the size-variable, it has been decided to sort the firms into three groups on book-to-market equity, and only into two groups on size. Further, the authors emphasize that when using the value-weighted returns on mimicking portfolios, the variance of factors is minimized and, what is even more important, the mimicking portfolios are more able to capture the very different behaviour of small/big (size) and high/low (value) stocks. The latter concern matters for displaying the realistic investment opportunities.

The dependent variable is also constructed as returns on the portfolios formed on size and book-to-market. Fama *et al* (1993) emphasize that it is important to form the dependent and explanatory portfolios using the same market information. That is, since the aim of the study has been stated to examine whether the mimicking portfolios associated with the underlying risk factors capture these common risk factors in equity returns.

The size and value factors are documented to be priced in the market (Fama *et al*, 1993), disregarding the fact that, at first glance, they are actually nonmarket risk factors. Moreover, they are documented to be important in explaining both the cross-sectional return distributions and the time series of returns. (Fama *et al*, 1993, 1995)

However, the debate on the topic of economic meaning of the Fama-French factors is still hot. There are several suggestions regarding interpretation of the size and value factors. Among them, there are empirical arguments for the size and value factors being proxies for leverage effects (Fergusson *et al*, 2003), and a statement that the value factor is a proxy for investor bias in earnings-growth extrapolation (Lakonishok *et al*, 1994).

It must be kept in mind that the three-factor model of Fama and French is basically one of the models that have been aimed to relaxing the strong “ideal-world” assumptions of CAPM, and that have provided evidence of the market portfolio not being a sufficient market risk proxy when used alone. (Chung *et al*, 2006) However, the experimental nature of the Fama and French model leaves a lot of questions concerning the employed factors unanswered.

2.2 Default Risk in Equity Returns: Prior Research

Quantification of the tradeoff between risk and expected return is one of the central problems of financial economics that has been addressed by the traditional CAPM and its modifications and extensions. These models focus on the market-related variables that are proxies for systematic risk factors in stock returns. The existing research interest for default risk factor as regards equity returns can be attributable to the fact that default risk is contemplated as a mixture of systematic and idiosyncratic risk components. (Chan-Lau, 2006)

However, the number of studies addressing the link between stock returns and default risk is still sparse. The first study that deals with this issue is probably a study by Rietz (1988). This author considers the excess equity returns as compensation demanded by investors for being exposed to extremely big losses provoked by an event like, say, economic recession. Since such event is normally followed by a chain of corporate failures, it can be argued that default risk contains a systematic risk component and may be an important determinant of stock returns development over time.

At the same time, there is empirical documentation that distress/default occurs mostly due to idiosyncratic risks, which leads to a conclusion of nonsystematic character of default risk. This is a result of the study by Altman (1993), for instance. The study documents that the bonds of deeply distressed companies earn lower than average subsequent returns. (Chan-Lau, 2006)

The more recent studies presented beneath, aim to revise the conventional asset pricing models by accounting for default risk beyond the “classical” risk factors. Despite many similarities, the studies report different and even opposing empirical results.

Vassalou *et al* (2004) provide a perspective for assessment of the effect of default risk on equity returns by testing whether default risk is systematic. The default risk measure is obtained from the structural Merton (1974) model for credit risk measurement, deducing the distance-to-default estimate (measured in standard deviations) from the structural relationship between equity, debt and asset value. Applying the portfolio formation procedure on the size and value factors as provided by Fama *et al* (1993) and employing the time series of aggregated default risk measure, the authors testify to default risk factor being priced in equity returns. Moreover, they conclude that the Fama-French factors, size and value, are proxies for default related information. The latter finding is provided by studying the statistical properties of the portfolios, and is confirmed by the results of regression analysis.

Gharghori *et al* (2007) have arrived at the empirical findings opposing Vassalou's *et al* (2004), despite the implementation of a very similar methodological framework. Both studies basically investigate whether the size and value factors capture the priced default risk. However, the study of Gharghori *et al* (2007) is performed on the data material of the Australian equity market, whereas a similar study of Vassalou *et al* (2004) uses the U.S. data.

Garlappi *et al* (2008) provide an explanation of cross-sectional properties of equity returns addressing a question of how leverage level may influence equity returns through default risk. The authors state that leverage affects the dynamics of equity returns differently from how it influences the dynamics of the firm's asset returns. Their study aims to control for relationship between stock return and default risk while extending the model of Fama *et al* (1993) by default risk factor and accounting for the potential shareholder recovery upon financial distress. The results of the study are very comprehensive and suggest the equity return being hump-shaped in default probability just due to shareholder recovery upon distress.

Anginer *et al* (2010) have presented the results that contradict both financial theory regarding "risk – return" relationship and some other empirical findings. These authors confirm neither the hump-shaped relationship between equity returns and default probability (Garlappi *et al*, 2008) nor abnormally high returns on distressed stocks due to investor compensation for additional risk bearing (Vassalou *et al*, 2005). Instead, anomalously low returns for distressed stocks have been documented by Anginer *et al* (2010). They have also examined different models for credit risk measurement and arrived at a conclusion that corporate bond spread as a proxy for default risk outperforms structural credit risk models, bond ratings, and accounting information models. They emphasize that corporate bond spread variable explicitly accounts for systematic component of distress risk and is therefore a good proxy for risk-adjusted probability of default. The main result of the study of Anginer *et al* (2010) is that the default risk is not priced in equity returns. However, a distressed stock behavior is shown to be dependent on a set of characteristics, also in some way related to default risk, such as leverage, volatility, and profitability. So, generally speaking, the results of this study are somewhat inconclusive.

It is noticeable that a significant portion of existing on this field studies revise the three-factor model of Fama *et al* (1993) by testing whether the Fama-French factors are associated with systematic portion of default risk. The large positive changes in the size factor may be linked to increase in systematic default risk. That is because small firms are more likely to default than big firms and so should offer higher returns. The value factor is associated with relative profitability and therefore can also be contemplated as default risk factor. This can be explained by the fact that lower returns on the assets of high book-to-market firms imply lower creditability compared to low book-to-market firms. (Chan-Lau, 2006)

2.3 Default Risk Measurement

The models for measuring credit risk or default probability can be sorted into several classes by their theoretical grounds. Thus, there can be distinguished structural models (Merton, 1974; Brockman *et al,* 2003; Vasicek, 1984), accounting-based models (Altman, 1968; Bunn, 2003), models considering macroeconomic information and default correlations (Bonfim, 2009; Gersbach *et al*, 2003), and also theoretical ideas of utilizing of different market proxies for default risk (Anginer *et al*, 2010). However, with respect to the objectives of this thesis, structural models are judged to be most appropriate.

2.3.1 Structural models

The Merton (1974) approach of measuring default risk is a classical method that has even lived through some modifications. The Merton model belongs to a family of so called structural models that utilize market information as input and is based on the idea of valuing risky bonds and loans relying on the option pricing theory. Thus, structural models are grown upon the Black-Scholes option pricing framework.

The company's face value of debt, market value of equity and market value of assets (estimated on the basis of the market value of firm's equity) are unified within the structural approach using the Black-Scholes option pricing terminology. In this theoretical setting the stockholders are said to own a call option on the firm's assets. Then, the strike price of this call option is equivalent to the debt level of the firm. Probability that the option will not be exercised is basically the probability of default. (Dionne *et al*, 2008)

Merton (1974) provides a pricing technique for a corporate security under condition of significant default probability. In essence, he studies corporate debt and risk structure of interest rate in connection to risk factors. Merton argues that both amount of corporate debt and interest rate depend on the following factors: risk free rate of return, contract conditions and default probability of the firm.

Further, Merton argues that "riskiness" *per se* is not much about interest rate of the economy; instead, it is connected to unanticipated changes in corporate default probability. (Merton, 1974)

The conclusions of Merton (1974) find the support and inspiration in the extension of the Black-Scholes setting for option pricing to methodological framework for pricing of corporate liabilities. Corporate liabilities in the Black-Scholes world are contemplated as options. (Black *et al*, 1973)

According to Black *et al* (1973), the bondholders of the firm, viewed as owners of the firm's assets, enable the stockholders to buy back the assets at the time of bonds maturity – that is, by writing an option to stockholders. At bonds maturity the common stock value is therefore a non-negative difference between the asset value and the face value of debt. This amount equals the value of the option which is a function of time and stock price. Consequently, the bond value is simply a difference between stock price and option value.

As stated above, this reasoning has been used by Merton (1974). Utilizing the option theory, Merton shows that at time of bond maturity the value of equity is positive if the asset value is greater than the debt face value; however, the equity value is zero if the asset value at time of bond maturity drops below the debt face value. In the latter case the default event occurs. Appealing to the option terminology, the two described conditions result into the following expression for the bond value: $F(V, 0) = min[V, B]$. Thus, the bond value is a function of the firm value and time to maturity and takes on the lowest value between firm value and debt face value. (Merton, 1974)

Merton (1974) emphasizes that there exists an isomorphic relationship between firm value and value of a particular corporate security. He illustrates this by, for instance, perfectly correlated returns on firm assets and a particular corporate security. Moreover, he states that a firm value and a particular corporate security value are affected by common variables – interest rate, business risk (volatility of firm assets), and current and future payout policy.

However, the structural model provided by Merton has become a subject for criticism for not considering the probability of defaulting before bonds maturity. A methodological response to this problem has been proposed by Brockman *et al* (2003) by introduction of a barrier structural model. This model contemplates equity as a down-and-out call option on the firm assets, and herewith attaches the option value over its entire life to the time-development of the underlying asset value.

In fact, the barrier structural modeling provides the bondholders with a down-and-in call option. This option may be activated before the total deterioration of the firm value, which in practice implies debt re-negotiations. (Brockman *et al*, 2003)

Vasicek (1984) has developed the Merton (1974) approach by combining it with contingent claim modeling (similarly to the barrier structural framework). Thus, a firm defaults as it reaches a predetermined barrier; so the distance-to-default measure is expressed in standard deviations of asset market values and shows by how many standard deviations the asset market value is away from the default point. (Saunders *et al*, 2002) The KMV™ model uses a huge database of expected default frequencies (EDF) developed on the basis of numerous observations of default points of defaulted firms in relation to their asset market values at time point of default. The structural model of Vasicek (1984) is built on three essential driving factors of default risk – market value of assets, asset volatility, and default point.

Asset market value is basically a value of the firm. It naturally depends on the equity market value, stock volatility, and liability structure. Asset market value can be seen as a present value of the future cash flows generated by the assets. (Bohn *et al*, 2003)

Asset volatility is attributable to business and industry risk and technically is an annual standard deviation of the asset market value. Higher volatility implies higher uncertainty about asset market value, and higher probability that the asset market value drops below the default point. (Bharath *et al*, 2004; Bohn *et al*, 2003)

Default point is a barrier that implies defaulting. Generally, the default event occurs when the value of the firm falls below this barrier. However, it must not be always a case, since a specific liability structure (such as large portion of long term debt) may allow renegotiation upon default. Then, the default point would be given by a difference between asset market value and company's debt, which is a negation of market value of equity. Therefore, the default point is a function of a particular firm's liability structure. (Bohn *et al*, 2003)

However, the structural models are subject to some critique because they can be applied only on the listed companies, since they involve equity price information. Another ground for criticism is so called trading noises which can negatively influence the accuracy of asset volatility estimates. (Bunn, 2003)

2.3.2 Some Important Remarks Regarding the Chosen Default Risk Measure

Estimates of distances-to-default defined structurally seem to be appropriate as proxy for default risk. That is because distance-to-default is a rather explicit measure of a firm's financial health which is mirrored in the changes in enterprise value, asset volatility and hence creditability – via deviations from the optimal capital structure. Translated into default probability (using the empirical database as by KMV™, or applying the assumption of normally distributed asset returns), this default risk measure is relatively easy to interpret.

On the other hand, the market proxies for default risk, referred to in the introductory chapter, are technically more simple and transparent, but their isomorphic connection to the creditability and default via changes in enterprise value is quite peculiar and may require an insight in every particular corporate case.

However, one must be aware of the following modeling peculiarities what regards structural defining of distance-to-default. There is a possibility that distance-to-default can be over- or underestimated. If a sudden decrease in leverage takes place, the asset volatility can be overestimated and so also default probability, which implies underestimation of distance-to-default. Conversely, if a sudden increase in leverage takes place, the asset volatility is likely to be underestimated, and hence default probability, which implies overestimation of distance-to-default. (Bohn *et al*, 2003)

Another aspect of defining default risk structurally corresponds to re-negotiating and re-structuring ability of a particular firm. Financial restructuring devices are a direct response to distress and default risk and may involve asset sales, cutting dividends, equity restructuring, and debt restructuring – all for cash reserves building and leverage decreasing. Therefore, for two different firms, that display the same distance-to-default, the actual threat of defaulting may be different due to potential possibilities of renegotiation with debtholders on the lender concessions, and/or due to ability of exploiting various devices of financial restructuring.

3 Methodology and Data Collection

This chapter explains the chosen approach – primarily for the dependent portfolios formation, mimicking portfolios modeling, factor risk premiums estimation and testing, and distance-to-default calculation. That is, alongside with generation of the working hypotheses for relationship "default risk – average stock returns" and concerning the size and value variables being proxies for default related information. At last, the data collection process is described.

3.1 Thesis Methodology and Approach

3.1.1 Dependent Variable: Portfolios Formation

The dependent variable is represented by the two sets of time series. The first set contains four time series of returns on four portfolios, respectively; these four portfolios are constructed on size and book-to-market equity and are used in the regressions that estimate the three-factor model (containing market risk factor, size factor, and value factor). The second set contains eight time series of returns on eight portfolios, respectively; these eight portfolios are constructed on size, book-to-market and distance-to-default and are used in the regressions that estimate the four-factor model (where the classical three-factor model is extended by default risk factor).

The portfolio formation technique is used for sorting the stock returns into different groups. A similar procedure can be found in the study by Fama *et al* (1993).

For the three-factor model estimation, each of the two data sets – the data set of market capitalization (size) and the data set of book-to-market (value) – have been split into two groups. That is achieved by finding the 50^{th} percentiles of the ordered data samples for size and value, separately. Further, the four portfolios are constructed from the intersections of the two size-groups and two value-groups. Thus, the following portfolios represent the firms with certain characteristics: "small size – low value", "small size – high value", "big size – low value", "big size – high value".

Next, the stock excess returns must be allocated to the portfolios in accordance with size and value characteristics of the corresponding firms, and the averages of returns on the portfolios must be calculated. In this way, the four time series of equally-weighted monthly excess stock returns on the portfolios with different corporate characteristics are obtained.

For the four-factor model the median split has been applied on the three data sets – the data set of size, the data set of value, and the data set of distance-to-default. That is achieved by finding the 50^{th} percentiles of the ordered data samples for size, value, and distance-to-default, separately. Further, the eight portfolios are constructed from the intersections of the two size-groups, two value-groups, and two distance-to-default-groups. Thus, the following portfolios represent the firms with certain characteristics: "small size – low value – short distance-to-default", "small size – high value – short distance-to-default", "big size – low value – short distance-to-default", "big size – high value – short distance-to-default", "small size – low value – long distance-to-default", "small size – high value – long distance-to-default", "big size – low value – long distance-to-default", "big size – high value – long distance-to-default".

Next, the stock returns must be allocated to the portfolios in accordance with size, value, and distance-to-default characteristics of the corresponding firms and the averages of returns on the portfolios must be calculated. In this way, the eight time series of equally-weighted monthly excess stock returns on the portfolios with different corporate characteristics are obtained.

3.1.2 Independent Variables: Mimicking Portfolios Formation

The independent variables in this study are presented by four different risk factors – market risk factor, size factor, value factor, and default risk factor. The market risk factor is common for the two sets of regressions, and is represented by a time series of the NASDAQ excess returns index. Monthly index returns have been deduced from the NASDAQ Composite price index, followed by the risk free rate subtraction.

In order to form the remaining factors, the technique for constructing factor mimicking portfolios is used, such that the factors are mimicking portfolios of stock returns. The purpose of this method is to mimic the underlying risk factors. This must be done separately for each of the regression sets such that both the dependent portfolios and the explanatory mimicking portfolios contain the same underlying information. That is, since it must be determined whether the mimicking portfolios capture the common risk factors in equity returns. If they do, the factors are systematic risk factors and are priced in the market.

The mimicking portfolios formation is based on the methodology of Fama *et al* (1993). However, the adopted by this study way to introduce the default risk factor into the asset pricing model is not described in the existing literature yet. It must be noticed that a testimony to the success of the mimicking portfolio formation is when the correlations between returns on the mimicking portfolios is rather low. (Fama *et al*, 1993)

For the three-factor model estimation, the two data sets – the data set of size and the data set of value – have been split into two and three groups, respectively. That is achieved by finding the 50^{th} percentile of the ordered data sample for size and the 30^{th} and 70^{th} percentiles of the ordered data sample for value.[2] Further, the six portfolios are constructed from the intersections of the two size-groups and three value-groups as described in the previous subsection. When the monthly stock excess returns have been allocated to the portfolios, the average returns on these portfolios must be calculated.

Then, the two factor mimicking portfolios can be defined on the basis of averaged returns for the portfolios with "small" and "big" (for size), and "high" and "low" (for value) characteristics. Thus, the size and value factors are represented by the equally-weighted monthly excess returns on the "small-minus-big" and "high-minus-low" mimicking portfolios, respectively. One can expect a negative relationship between size factor and average excess returns and a positive relationship between value factor and average excess returns.[3]

For the four-factor model estimation, the three data sets – the data set of size, the data set of value, and the data set of distance-to-default – have been considered. The data set of size and the data set of value have been split into two groups each, by finding the 50^{th} percentiles of the ordered data sample of size and value, separately. The data set of distance-to-default has been split into three groups by finding the 30^{th} and 70^{th} percentiles of the ordered data sample for distance-to-default. The applied splits seem to be reasonable since it is expected that the default

---

[2] This specification follows the study by Fama *et al* (1993) and is argued for in the section *2.1 The Three-Factor Model by Fama and French (1993)*.

[3] In accordance with reasoning in the section *2.1 The Three-Factor Model by Fama and French (1993).*

risk factor may exhibit stronger properties of common risk factor in average returns, in comparison to the remaining candidate risk factors.

Further, the twelve portfolios are constructed from the intersections of the two size-groups, two value-groups, and three distance-to-default-groups as described in the previous subsection. When the monthly stock excess returns have been allocated to the portfolios, the average returns on these portfolios must be calculated.

Then, the three factor mimicking portfolios can be defined on the basis of averaged returns on the portfolios with “small” and “big” (for size), “high” and “low” (for value), and “short” and “long” (for distance-to-default) characteristics. Thus, the size, value, and default risk factors are represented by the equally-weighted monthly excess returns on the “small-minus-big”, “high-minus-low”, and “short-minus-long” mimicking portfolios, respectively.

As stated earlier, one may expect a negative relationship between size factor and average excess returns and a positive relationship between value factor and average excess returns. However, the relationship between default risk factor and average excess returns does not seem to be unambiguous and deserves to be discussed in a separate subsection.

3.1.3 The Working Hypothesis Regarding Relationship “Default Risk – Average Stock Return”

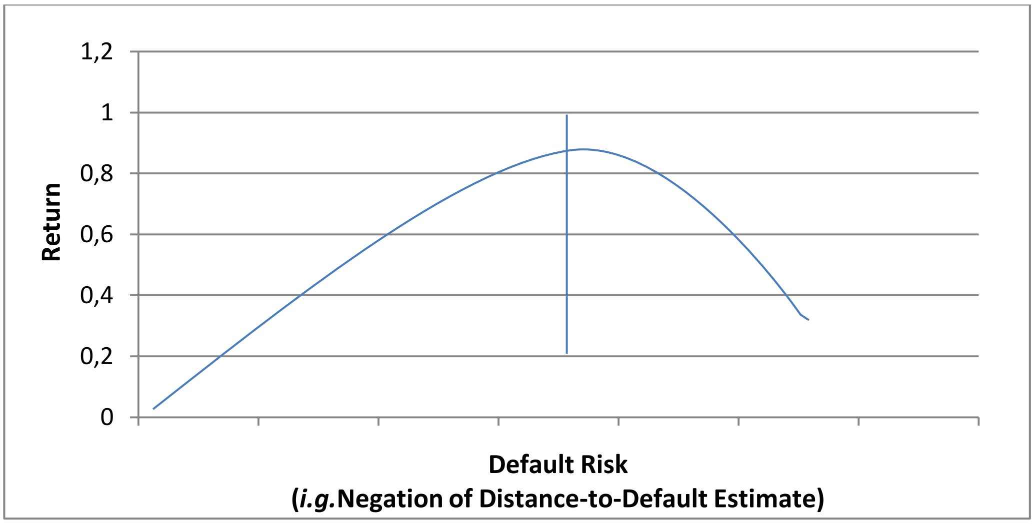
Default risk can be contemplated in two different contexts. On the one hand, a standard understanding of relationship between risk and return rests on the statement “higher risk – higher expected return”. In this case, a shorter distance-to-default is associated with a greater risk exposure, and hence higher expected return; thus, a negative relationship between distance-to-default and average returns is to expect.

On the other hand, a positive relationship between distance-to-default and stock return is also possible. This can be explained by negative market reaction on the event of corporate distress or just creditability worsening. The supporting empirical evidence for both statements can be found in the existing literature on this topic. (For comparison see *e.g.* Vassalou *et al*, 2004 and Gharghori *et al*, 2007)

Distance to the default point *per se* seems to be a reasonable explanation for the differences in empirical findings regarding the "default risk – average stock return" relationship. A relatively long (and also time-persistent) distance to the default point is naturally associated with a relatively strong "financial health" of the firm. Then, the undesirable variations in asset value and volatility associated with changes in financial stability are less likely to be perceived as events leading to a threat of distress and default. In this case, a required rate of return is expected to rise. However, when a distance to the default point is persistently short, its eventual additional shortening may provoke a negative market reaction, and hence a decrease in stock return is to expect.

On the diagram below the firms of the first type are allocated to the left of the vertical line, whereas the firms of the second type are allocated to the right of the vertical line.

Figure 1. Stock Returns as a Function of Default Risk



The prominent feature of the firms in the working sample is that their distances-to-default are rather long (however, the time series of distances-to-default in question exhibits a sufficient for statistical inferences variation). That is why it seems to be reasonable to expect a negative relationship between distance-to-default and average returns, hence positive relationship between default risk and average stock return.

On these grounds the returns on mimicking portfolio for the default risk factor equal to the difference between equally-weighted monthly excess returns on a “short distance-to-default” portfolio and equally-weighted monthly excess returns on a “long distance-to-default” portfolio. The expected relationship between average stock returns and default risk factor mimicked by the “short-minus-long” portfolio is positive.

3.1.4 Factor Mimicking Portfolios Formation with Respect to Market Inefficiency

It is of importance to form mimicking portfolios in an optimal way, allowing the market to react on the changes in corporate information. That is why a description of portfolio formation process with respect to prevailing market inefficiency is presented below. The study by Fama *et al* (1993) is followed when constructing the size and value factors. However, there are no studies that could suggest on construction of default risk factor.

In order to construct the size factor mimicking portfolio, the market capitalization variable has been lagged one month. Thus, the average monthly excess returns on this mimicking portfolio may be seen as a result of a one-month-delayed response of the market on the information on market capitalization.

What regards construction of the value factor mimicking portfolio, the average monthly returns on this portfolio is basically a market response on the book-to-market equity information that had become available13 months ago. To be more precise, when calculating the book-to-market ratio, a 13-months-lagged book value and a 7-months-lagged market value have been taken. In other respects, the book equity has been used as a sum of common equity and deferred taxes; the firms with negative book equity values have been excluded from the working sample.

Further, the average monthly excess returns on the default risk factor mimicking portfolio are decided to be a result of a one-month-delayed response of the market on the information underlying default risk.

3.1.5 Factor Risk Premium: Estimation Technique, Empirical Tests, and Modeling Specifics

The empirical modeling applied to this study is primarily premised on the three-factor model developed by Fama *et al* (1993) and enhanced by the default risk factor, as stated earlier. The experimental design for both the three- and four-factor model is the same, differing just by the number of factors, and implies, beyond parameters estimation, evaluation of the factor risk premiums. This section discusses the factor risk premium estimation referring only to the four-factor model due to the mentioned similarities.

One of the objectives of this thesis assignment is to determine whether the risk factors are associated with systematic risk sources and therefore are priced in the market. If a risk factor is common in equity returns, it is considered being systematic. This means that this risk is shared by all (or by the majority of) financial actors, and therefore is undiversifiable. (Fama *et al*, 1993; Campbell *et al*, 1997) In this case it is reasonable that an investor expects a risk premium for the systematic risk exposure. Thus, in empirical terms, it can be stated that the risk exposure is priced in the market if the corresponding factor risk premium is positive (which must not hold for hedge factors, however) and statistically different from zero. (Campbell *et al*, 1997)

The exact factor pricing model for expected returns on traded assets should be written as follows:

$$\mu = \iota\lambda_0 + \mathrm{B}\lambda_{\mathrm{K}},$$

where B is the factor sensitivity matrix, λ_0 is the riskfree rate or the zero-beta expected return, λ_{K} is the factor risk premium. (Campbell *et al*, 1997) Since the current study exploits the excess returns on the mimicking portfolios of the underlying risk factors, the factor risk premiums can be estimated directly from the sample means of the excess returns on the portfolios (Fama *et al*, 1993; Campbell *et al*, 1997):

$$\lambda_K = \frac{1}{T}\sum_{t=1}^{T} F_{Kt}.$$

The risk factors in this study have been specified on the theoretical grounds; therefore it is of interest to see if they are priced in the market. This may be achieved by testing whether the estimated risk premiums are significantly different from zero by using the conventional test statistic:

$$\varphi_k = \frac{\hat{\lambda}_K}{\sqrt{Var(\hat{\lambda}_{KK})}} \sim \mathcal{N}(0,1),$$

where $\hat{\lambda}_K$ is the estimated factor risk premium, and $Var(\hat{\lambda}_{KK})$ is the *KK*th element in the sample variance-covariance matrix of the risk factors. (Campbell *et al*, 1997) The estimator of the sample variance-covariance matrix of the risk factors is obtained from the $T \times K$ matrix of the risk factors according to the following expression:

$$\widehat{Var}[\hat{\lambda}] = \frac{1}{T}\hat{\Omega}_K = \frac{1}{T^2}\sum_{t=1}^{T}(F_t - \frac{1}{T}\sum_{t=1}^{T}F_{Kt})(F_t - \frac{1}{T}\sum_{t=1}^{T}F_{Kt})'.$$

The cumulative standard normal distribution is applied on the value of the test statistic for the critical value obtaining.

Additionally, in order to test whether the risk factors are jointly priced, the following test statistic of the null hypothesis that the factors are jointly not priced can be applied:

$$\varphi = \frac{(T-K)}{TK}\hat{\lambda}'_K \widehat{Var}[\hat{\lambda}_K]^{-1}\hat{\lambda}_K,$$

where $\hat{\lambda}_K$ is the $K \times 1$ vector of estimated factor risk premiums, and $\widehat{Var}[\hat{\lambda}_K]^{-1}$ is the inverse of the sample variance-covariance matrix of the risk factors. The given test statistic is, under the null hypothesis, asymptotically *F*-distributed with *K* and *T–K* degrees of freedom, which determines the critical value for the test of interest. (Campbell *et al*, 1997)

At the same time, it is possible to contemplate the estimation of the factor risk premiums and regression parameters of the risk factors by a single experimental model presented by the system of following equations:

$$\left\{\begin{array}{c} R_{it} = b_i R_{mt} + s_i SmB_t + v_i HmL_t + d_i SmL_t + \varepsilon_{it} \\ R_{mt} = \lambda_m + \varepsilon_{mt} \\ SmB_t = \lambda_{SmB} + \varepsilon_{st} \\ HmL_t = \lambda_{HmL} + \varepsilon_{vt} \\ SmL_t = \lambda_{SmL} + \varepsilon_{dt} \end{array}\right\},$$

where R_{it} and R_{mt} are excess returns on the dependent portfolios and the market portfolio, respectively; SmB_t, HmL_t, and SmL_t are the excess returns on the factor mimicking portfolios of the underlying size ("small-minus-big"), value ("high-minus-low") and default risk ("short-minus-long"), respectively; λ_m, λ_{SmB}, λ_{HmL}, and λ_{SmL} are the factor risk premiums of the market risk factor, size factor, value factor and default risk factor, respectively.

In essence, the four latter equations, containing the factors risk premiums, are namely the mean-adjusted transformations of the explanatory variables (market risk factor and mimicking portfolios of the underlying size, value and default risk factors). This can be seen as an alternative way to illustrate the empirical nature of factor risk premiums.

Moreover, contemplation of the phenomenon of factor risk premium can lead to the discussion on mispricing. If asset returns can entirely be explained by empirical model depicting the relationship between asset returns and factor risk premiums via factor loadings (Campbell *et al*, 1997), the mispricing indicator would be expected being equivalent to the following expression: $\alpha^{*}=b_i\lambda_m + s_i\lambda_{SmB} + v_i\lambda_{HmL} + d_i\lambda_{SmL}$ (Gharghori *et al*, 2007).

3.1.6 The Working Hypothesis Regarding the Size and Value Factors being Proxies for Default Related Information

If default risk is priced in equity returns, then the corresponding risk factor exhibits a positive and statistically different from zero factor risk premium. The same holds for the factors size and value which are already shown (Fama *et al*, 1993; Vassalou *et al*, 2004) to be related to systematic risks sources. It has also been shown that the size and value factors are proxies for some default related information. (Vassalou *et al*, 2004)

It is plausible to hypothesize that the size and value factors can be seen as proxies for systematic portion of default related information, if these variables lose their properties of risk factor sensitivities for stocks as soon as the three-factor model has been augmented with default risk factor. Empirically, it implies that if the estimated coefficients of the size and value variables appear to be significant within the three-factor model, but lose their explanatory power what regards equity returns within the four-factor model, then the size and value factors can be seen as proxies for the systematic component of default risk.

The Fama-French factors are expected to be associated with the systematic portion of default risk, if and only if the default risk factor appears to be common in equity returns and hence priced in the market.

3.1.7 An Anatomy of Defaulting in a Structural Setting: Calculation of Distance-to-Default

Distance-to-default calculation procedure used in this thesis assignment follows the main principles of structural modeling. In other words, establishment of the structural relationship between market value of equity, face value of debt and market value of assets allows for solving for distance-to-default as default risk measure.

However, the market value of assets and asset volatility are not directly observable and must be extracted from the market value of equity and equity volatility, respectively. The Black-Scholes option pricing framework provides a formula that depicts the relationship between market value of equity and market value of assets:

$$V_E = V_A N(d_1) - e^{-r_f T} X N(d_2),$$

where V_E corresponds to the market value of equity, V_A is the market value of assets, X is the face value of debt, T is the time horizon, and r_f is the risk free rate. In this equation $N(\cdot)$ is the cumulative normal distribution, $d_1 = \frac{ln\left({V_A}/{X}\right)+\left(r_f+\frac{1}{2}\sigma_A^2\right)T}{\sigma_A\sqrt{T}}$, and d_2=$d_1 - \sigma_A\sqrt{T}$, where σ_A is the volatility of asset market value. (Bohn *et al*, 2003) Thus, the firm's equity value is expressed as a function of the value of firm's assets. This is the first important equation used in the distance-to-default model.

The second important equation the distance-to-default model rests on, must depict a relationship between equity volatility and asset volatility. According to the assumption that equity is a function of asset value and by applying Itô's Lemma, $\sigma_E = \left(\frac{V_A}{V_E}\right)\frac{\partial V_E}{\partial V_A}\sigma_A$, where σ_E is the equity volatility, and the other parameters are defined as above. (Bharath *et al*, 2008) In the continuous time setting it can be shown that $\frac{\partial V_E}{\partial V_A} = N(d_1)$, and thus $\sigma_E = \left(\frac{V_A}{V_E}\right) N(d_1)\sigma_A$. (Pennacchi, 2008) For a better intuitive understanding, it can be written even in a following way: $\sigma_E = \frac{V_A}{V_E}\Delta\sigma_A$. (Bohn *et al*, 2003)

Thus, the system of non-linear equations must be solved simultaneously and the parameters V_A and σ_A must be inferred under assumption of normally distributed asset returns.[4] The market value of equity is directly observable since the companies included in the study are listed on the stock exchange; equity volatility of the firms in question is calculated on the basis of yearly equity returns; the risk free rate used in the model is a one-year U.S. Treasury bill; the time horizon is set to one year implying that the obtained values of distances-to-default correspond to default probabilities that a particular firm defaults within a one-year period; the default point is defined as a sum of the face value of short-term debt and a half of the long-term liabilities, which is a commonly used way. (Bharath *et al*, 2008; Chan-Lau *et al*, 2006; Korablev *et al*, 2007)

As soon as the asset market values and the values of volatility of assets have been obtained, the yearly distances-to-default can be easily calculated using the formula:

$$DD = \frac{\ln\left({}^{V_A}/_{V_E}\right) + \left(\mu + \frac{1}{2}\sigma_A^2\right)T}{\sigma_A\sqrt{T}},$$

where μ is the expected asset return per time unit. Moody's KMV™ retrieves the asset returns from the initial asset volatility: the values of assets returns in the current period are used for determining the next period's asset returns with respect to the corresponding asset values. The process continues until convergence. (Bohn *et al*, 2003)

Indeed, it is hard to calculate the expected asset returns; therefore, different proxies can be used. Exploiting the fact that the changes in asset value must be closely related to stock returns, the expected asset returns might be computed as means of changes in asset value: $\mu = \frac{1}{T}\sum_{t=1}^{T} \Delta ln(V_A)$. This finds support in the study of Da *et al* (2006), where asset returns (or asset value drift) are shown to be related to the drift in equity returns through the sensitivity of equity value with respect to asset value: $\mu_E - r = \frac{\partial V_E}{\partial V_A} \cdot \frac{V_A}{V_E}(\mu - r)$.

At the same time, Saunders *et al* (2002) suggest using a simplified and intuitively understandable version of the formula for distance-to-default:

$$DD = \frac{V_A - X}{V_A \sigma_A}.$$

---

[4] The comments on MathCAD utilization when solving the specified system of equations, the details about used iteration methods, the modifications applied on the input and output data are presented in *Appendix II* (sections *II.1* through *II.4*)

The latter expression provides an almost visual interpretation of the default measure as a distance between asset market value and default point, expressed in standard deviations. The similarity of the results given by the both formulas might be expected.

3.2 Data Collection and Information Source

The current study investigates the non-financial companies listed on the NASDAQ stock exchange. Due to the time constraints, the entire population of the NASDAQ equities (which is represented by more than 2700 listed companies) is not included in the working sample. Instead, after the financial firms, such as banks, investment and insurance companies, have been sorted away, alongside with companies reporting negative values for book equity, the working sample is formed by 171 randomly selected companies.

It must be noticed that among these randomly selected firms there are also those with incomplete history of financial reporting. Since the model estimation process in this study uses averaged portfolio returns as regression inputs, some absent from the financial statements information concerning some few firms does not influence the models' estimation capabilities. Conversely, the presence of the firms with incomplete history of financial reporting ensures that there is no selection bias implied by possible exclusion of the firms actually defaulted during the estimation period.

The firm specific data used in this thesis consists of the time series of share prices observed on a monthly and yearly basis, yearly observations for the short- and long-term debt, the time series of yearly observations for common equity and deferred taxes (for the book equity calculation), and the market capitalization observed on a yearly basis. The market specific data consists of the time series of monthly and yearly returns on a one-year U.S. Treasury Bill and the NASDAQ Composite price index observed on a monthly basis.

The length of the collected time series is 19 years and one month (February 1991 – February 2010). The estimates of distances-to-default are obtained on a yearly basis; the regression analysis is conducted on a monthly basis. The sample length is considered to be sufficient for a time series analysis.

All the time series have been collected from the DataStream financial database. It must be kept in mind that the DataStream database is a secondary source of corporate and market information which reserves the right to imperfections in the data material.

4 Empirical Findings

This chapter reports the empirical results of the study. It naturally starts by inspection of the graphical patterns of the time series involved. Then, the descriptive statistics of the data material is presented and discussed. The regression estimation results are followed by the assessment of the models' explanatory properties. At last, the evaluation results of the factor risk premiums are reported.

Before reporting the empirical results of the study, some introductory words may be needed in order to emphasize a specificity of the regressions' input data. This thesis assignment deals with the time series of returns on portfolios constructed in accordance with technique described in detail in the methodology chapter. The study does not use the time series directly observable in the market. Moreover, the empirical design of the study is based on a descriptive-normative approach, when the research goals are focused on the explanation of potential relationships between the variables in question rather than on determining the optimal models. These facts allow not focusing too much on the diagnostic testing of the input time series. However, it is accounted for some time series characteristics such as normality criterions and correlations between explanatory returns.

4.1 The Results of an Ocular Inspection of the Time Series of Returns

Analyzing the graphs of the time series of returns, one can observe a significant jump approximately in the middle of year 2008.[5] This fact documents the evidence of the financial crisis that the entire world entered in year 2007 and which became apparent in the U.S. with the bankruptcy of Lehman Brothers.

[5] The graphs of the time series based on the original data matherial can be found in *Appendix III* (*Figure III.1*)

As such, for statistical quality's sake, it has been decided to exclude the latter observations in the time series, starting from March 2008 when the structural alteration has been originated.[6,7] In this particular case the exclusion of a fragment of the time series is preferable to structural break modeling. This is due to the fact that the financial crisis is not over at the time when the study is conducted; and so, it is unclear how the time series of market and corporate variables are going to develop. For comparison, the structural break modeling is appropriate when different market conditions are established after the break-point has occurred.

4.2 Descriptive Statistics of the Variables in Question

The descriptive statistics for each time series, used in the regression analysis, is of interest. It is expected that when the input time series exhibit the distribution characteristics close to normality, the models they are involved in, generate the most reliable parameters estimates (significant and theoretically and empirically reasonable coefficients), relatively high R-squared values, and white noise residuals. Moreover, normality is one of the essential assumptions of the OLS-estimator and basically makes the translation of statistical inferences into the description of properties of the examined population (and possibly theory-generation) more comprehensible and transparent.

This section reports for descriptive statistics regarding the two sets of time series, each of which corresponds to one of the competing three- and four-factor models. As explicated in the methodology chapter, the input time series are different for the two asset pricing models and must be investigated separately. That is because the input time series are the time series of returns on the portfolios that aim to capture corporate and market information relatively to each particular asset pricing specification.

The results of the descriptive statistic tests provide an additional support for the necessity to exclude the observations starting from March 2008. Thus, what regards the time series involved in the three-factor model estimation[8], the dispersion of observations is minimized for

[6] Thus, the 22 latter monthly observations are excluded from the original working sample
[7] The graphs of the time series based on the sub-sampled data matherial can be found in *Appendix III* (*Figure III.2*)
[8] The results of the normality tests (Jarque-Bera statistic, skewness and kurtosis) together with the mean, median, maximum and minimum values and standard deviations are reported in *Appendix III (tables III.1 and III.2)* for both the original and reduced time series involved in the three-factor model estimation

the size (change in maximum value from 1,0990 to 0,1562 in relation to unchanged minimum value) and value (change in maximum value from 1,6521 to 0,7646 in relation to unchanged minimum value) variables. The same is observable for the dependent variable “small size – high value”. It must be noticed that a wide dispersion is definitely not anything negative in a statistical context. However, in the current case, the dispersion is not even, but caused by extreme outliers.

Still, the results of the Jarque-Bera tests reveal the fact of non-normally distributed returns for the majority of series of the sub-sampled data. The attempts to exclude the remaining outliers from the sub-sampled data have not led to any improvements of normality criterions. This can be explained by the sample size insufficiency which is a reason for relatively small portfolios of dependent returns and hence insufficient representation of the corporate characteristics that are meant to be captured by the portfolios in question.[9]

The summarized descriptive statistic testing results for the time series of returns involved in the four-factor model estimation provide similar distribution characteristics as the descriptive statistic tests performed on the time series involved in the three-factor model estimation.[10] However, the exclusion of the 22 latter monthly time series observations has led to mitigation of leptokurtic character of distribution (*e.g.* for the returns on the dependent portfolio “small size – high value – short distance-to-default” and for all three time series of returns on the factor mimicking portfolios).

Additionally, the correlations between the time series of returns on the explanatory portfolios used in the corresponding three- and four-factor models are found to be relatively low and are deemed to be satisfactory both in the portfolio formation context and for the regression analysis. The correlation matrices of variables of interest generated by the sub-sampled data matherial can be found in *Table 4.2.a* and *Table 4.2.b* beneath.

[9] This issue is discussed in the section *1.5 Delimitations and Weaknesses of the Study*
[10] The results of the normality tests (Jarque-Bera statistic, skewness and kurtosis) together with the mean, median, maximum and minimum values and standard deviations are reported in *Appendix III (tables III.3 and III.4)* for both the original and reduced time series involved in the four-factor model estimation

Table 4.2.a. Correlations between the Time Series of Returns on the Explanatory Portfolios Involved in the Three-Factor Model Estimation – Sub-sample

	RM	SIZE	VALUE
RM	**1**	-0,2592	-0,1743
SIZE	-0,2592	**1**	-0,6250
VALUE	-0,1743	-0,6250	**1**

Table 4.2.b Correlations between the Time Series of Returns on the Explanatory Portfolios Involved in the Four-Factor Model Estimation – Sub-sample

	RM	SIZE	VALUE	Distance-to-Default
RM	**1**	-0,2896	-0,1574	0,2259
SIZE	-0,2896	**1**	-0,4310	-0,3845
VALUE	-0,1574	-0,4310	**1**	0,3980
Distance-to-Default	0,2259	-0,3845	0,3980	**1**

4.3 Results of the Regression Estimations

In this section the regression estimation results generated by the two asset pricing models of interest are presented. Four time series regressions for the three-factor model and eight time series regressions for the four-factor model are estimated by OLS using both the entire and reduced data samples. As motivated earlier, the working sample is reduced by exclusion of the 22 latter monthly time series observations. Therefore, the regression estimation results shown in this section are based on the sub-sampled data matherial.[11]

The three-factor model estimation results are summarized in *Table 4.3.a*. The estimated coefficients of all explanatory variables are found to be significant in the four regressions that correspond to the four dependent portfolios. Additionally, the *R*-squared values for all four regressions are fairly high. However, the regression model for the "big size – high value" (BH) portfolio is the only one that provides the factor loadings with theoretically expected signs.

[11] The tables summarizing for the regression estimation results for the entire sample are placed in *Appendix III* (*tables III.5* and *III.6*) for comparison

Table 4.3.a. Three-Factor Model Estimation Results – Sub-Sample

	SL	SH	BL	BH
Intercept	**0,000**	**-0,003**	**-0,001**	**-0,006**
Std. Error	0,004	0,002	0,003	0,005
t-statistic	0,092	-1,418	-0,557	-1,395
Market risk factor	**0,784**	**0,741**	**0,749**	**0,799**
Std. Error	0,054	0,035	0,040	0,068
t-statistic	14,448	21,295	18,852	11,664
Size factor	**0,554**	**0,647**	**-0,361**	**-0,412**
Std. Error	0,078	0,050	0,057	0,099
t-statistic	7,068	12,859	-6,292	-4,159
Value factor	**0,190**	**0,461**	**-0,223**	**0,669**
Std. Error	0,053	0,034	0,039	0,067
t-statistic	3,557	13,465	-5,716	9,924
R^2	**0,530**	**0,702**	**0,777**	**0,715**
Sample length	**205 observations**			

Table 4.3.a reports the estimation results of the three-factor model, where the returns on portfolios "small size – low value (SL), "small size – high value" (SH), "big size – low value" (BL), "big size – high value" (BH) are regressed on the Market risk factor, Size factor and Value factor. The estimation period lies between February 1991 and February 2008. The table presents the estimated coefficients of the parameters for each regression alongside with the standard errors, t-statistics and the values of *R*–squared.

Table 4.3.b. Four-Factor Model Estimation Results – Sub-Sample

	SLS	SLL	SHS	SHL	BLS	BLL	BHS	BHL
Intercept	**-0,008**	**-0,001**	**0,002**	**-0,011**	**0,000**	**-0,004**	**-0,005**	**0,001**
Std. Error	0,004	0,004	0,003	0,003	0,004	0,002	0,005	0,004
t-statistic	-1,811	-0,190	0,624	-3,687	0,082	-1,456	-1,163	0,305
Market risk factor	**0,612**	**0,733**	**0,863**	**0,542**	**0,658**	**0,798**	**0,704**	**0,532**
Std. Error	0,067	0,062	0,046	0,045	0,056	0,037	0,069	0,055
t-statistic	9,115	11,901	18,704	12,046	11,679	21,418	10,210	9,594
Size factor	**0,700**	**0,426**	**0,875**	**0,409**	**-0,411**	**-0,153**	**-0,596**	**-0,061**
Std. Error	0,093	0,085	0,064	0,062	0,078	0,051	0,095	0,077
t-statistic	7,548	5,012	13,743	6,589	-5,286	-2,976	-6,264	-0,803
Value factor	**-0,277**	**0,202**	**0,645**	**0,414**	**-0,732**	**-0,101**	**1,268**	**0,310**
Std. Error	0,095	0,087	0,065	0,064	0,080	0,053	0,098	0,079
t-statistic	-2,918	2,311	9,878	6,502	-9,171	-1,914	12,994	3,946
Default risk factor	**0,495**	**0,022**	**0,175**	**0,049**	**0,325**	**-0,026**	**0,931**	**-0,219**
Std. Error	0,064	0,059	0,044	0,043	0,054	0,035	0,066	0,053
t-statistic	7,757	0,372	3,999	1,148	6,056	-0,720	14,200	-4,143
R^2	**0,530**	**0,438**	**0,702**	**0,465**	**0,691**	**0,770**	**0,860**	**0,365**
Sample length	**205 observations**							

Table 4.3.b reports the estimation results of the four-factor model, where the returns on portfolios "small size – low value – short distance-to-default" (SLS), "small size – low value – long distance-to-default" (SLL), "small size – high value – short distance-to-default" (SHS), "small size – high value – long distance-to-default" (SHL), "big size – low value – short distance-to-default" (BLS), "big size – low value – long distance-to-default" (BLL), "big size – high value – short distance-to-default" (BHS), "big size – high value – long distance-to-default" (BHL) are regressed on the Market risk factor, Size factor, Value factor, and Default risk factor. The estimation period lies between February 1991 and February 2008. The table presents the estimated coefficients of the parameters for each regression alongside with the standard errors, t-statistics and the values of *R*–squared.

Similarly, *Table 4.3.b* provides evidence of significant and theoretically sound coefficients obtained when estimating the four-factor regression model that aims to explain the returns on the “big size – high value – short distance-to-default” (BHS) portfolio. Besides, this regression model exhibits the highest in its class goodness-of-fit measure.

Also, the estimated coefficients for the intercepts obtained when estimating all the three-factor regressions and seven of eight[12] four-factor regressions, are found to be statistically zero. Very small and insignificant intercepts imply exact pricing.

However, as noticed above, the estimation results of only one regression model[13] within the set of three-factor regressions and only one regression model[14] within the set of four-factor regressions provide the estimated parameters that are theoretically sound. It is worth recalling that the modeling quality and statistical reliability in this study to a great extent depend on the variation within the formed portfolios. Analysis of the corporate characteristics used as grounds for returns allocation shows that relatively big (and consequently with a wider return spectrum) portfolios are exactly “big size – high value” and “small size – high value – short distance-to-default” dependent portfolios.

4.4 Explanatory Properties of the Regression Models

As one can recall, an ocular inspection of the graphs of the input time series has led to the decision to proceed working with a reduced data sample, excluding the 22 latter monthly time series observations. It has also been stated that the regressions built on the time series characterized by the most acceptable descriptive statistics are expected to generate the most statistically reliable parameters estimates. Moreover, the models which contain such time series have been expected to exhibit high explanatory power and originate the residual estimates that possess white noise properties.

[12] Except for the model aiming to explain the returns on the “small size – high value – long distance-to-default” portfolio
[13] That aims to explain the returns on the “big size – high value” portfolio
[14] That aims to explain the returns on the “small size – high value – short distance-to-default” portfolio

It has been shown that the reduction of the working sample has led to some negligible[15] improvements of descriptive statistics of the time series of interest, such as exclusion of extreme outliers and alleviation of the leptokurtic distribution manifestation.

It has also been noticed that the theoretically sound parameters estimates have been obtained by estimating the three-factor model explaining the returns on the "big size – high value" portfolio and the four-factor model explaining the returns on the "big size – high value – short distance-to-default" portfolio. However, it is believed to be so not due to the mentioned changes in the descriptive statistics of the input data, but rather due to the comparatively wider spectrum of returns within the portfolios involved, in other words simply because of the comparatively bigger portfolios.

Now, it can be of interest to examine whether the models that aim to explain the returns on the "big size – high value" and "big size – high value – short distance-to-default" portfolios are statistically better than others. Indeed, the four-factor model for the returns on the "big size – high value – short distance-to-default" portfolio exhibits the highest in its class explanatory power. However, it is not the fact for the three-factor model that explains the returns on the "big size – high value" portfolio.

Further, it is worth checking whether the time series of estimated residuals originated by the regression estimations are white noise. White noise residuals would indicate that the dependent returns are fully explained by the models employed. For this purpose the Breusch-Godfrey serial correlation test is used.

The testing results[16] provide evidence of presence of serial correlation in the four time series of estimated residuals originated by the four regression estimations of the three-factor model. What regards the four-factor regression estimations, the four of eight time series of the estimated residuals exhibit serial correlation. It must be emphasized that the estimated residuals originated by the four-factor model that explains the returns on the "big size – high value – short distance-to-default" portfolio are found to be possessing white noise properties.

---

[15] The improvements have appeared to be negligible due to the reasons discussed in the section *.1.5Delimitations and Weaknesses of the Study* and related to the portfolio formation specificity

[16] The results of the serial correlation tests are summarized in *Table 4.4.a* and *Table 4.4.b* for the estimated residuals of the three- and four-factor models, respectively

Table 4.4.a Result Summary of the Breusch-Godfrey Test – the Three-Factor Model – Sub-Sample

Breusch-Godfrey Serial Correlation LM Test	SL	SH	BH	BL
Prob. F(2,198)	0.000010	0.000000	0.000000	0.000000
Prob. Chi-Square(2)	0.000014	0.000000	0.000000	0.000000

Table 4.4.a summarizes the results for the Serial correlation LM test performed on the estimates of the residuals originated by the three-factor model estimations. Abbreviations SL, SH, BH, and BL stay for the "small size – low value", "small size – high value", "big size – high value", and "big size – low value" dependent portfolios, respectively, and are aimed to the series of estimated residuals originated by the corresponding models estimations. The estimation period lies between February 1991 and February 2008.

Table 4.4.b Result Summary of the Breusch-Godfrey Test – the Four-factor model – Sub-Sample

Breusch-Godfrey Serial Correlation LM Test	SLS	SLL	SHS	SHL	BLS	BLL	BHS	BHL
Prob. F(2,198)	0,01296	0,51115	0,00136	0,00000	0,00001	0,00017	0,11326	0,14425
Prob. Chi-Square(2)	0,01225	0,50034	0,00134	0,00000	0,00002	0,00018	0,10748	0,13732

Table 4.4.b summarizes the results for the Serial correlation LM test performed on the estimates of the residuals originated by the three-factor model estimations. Abbreviations SLS, SLL, SHS, SHL, BLS, BLL, BHS, and BHL stay for the "small size – low value – short distance-to-default", "small size – low value – long distance-to-default", "small size – high value – short distance-to-default", "small size – high value – long distance-to-default", "big size – low value – short distance-to-default", "big size – low value – long distance-to-default", "big size – high value – short distance-to-default", "big size – high value – long distance-to-default" dependent portfolios, respectively, and are aimed to the series of estimated residuals originated by the corresponding models estimations. The estimation period lies between February 1991 and February 2008.

Summarizing for this section, it can be stated that the four-factor models tend to outperform the three-factor models on average. The four-factor model that aims to explain the returns on the "big size – high value – short distance-to-default" portfolio[17], demonstrates (in its class of models) the highest explanatory power both in terms of goodness-of-fit measure and due to origination of white noise residuals.

[17] This model is referred to as "selected model" here-on

4.5 Factor Risk Premiums Estimation Results

Evaluation of the factor risk premiums[18] indicates whether the risk factors are common in equity returns and hereby systematic and priced in the market. In this section, the risk premium estimates and the corresponding test statistics are presented. *Table 4.5.a* summarizes for the single significance testing performed on the factor risk premiums attributable to the risk factors involved in the three-factor model.

Table 4.5.a Factor Risk Premiums: Single Tests for the Three-Factor Model's Factors

	Market risk factor	Size factor	Value factor
Factor risk premium	-0,027263106	0,002034324	0,010640424
Standard error of estimate	2,57382E-05	3,84426E-05	8,41806E-05
Test statistic	-5,373857457	0,328105914	1,159719189
Critical value	3,85349E-08	0,628584209	0,876918423

As can be seen, the risk premiums for the size and value factors are positive and demonstrate the standard errors of their estimates being very close to zero. However, only the value factor risk premium is significantly different from zero. Therefore, within the three-factor model, the risk underlying the value factor is common in equity returns and hence is considered to be priced in the market.

Testing joint significance of the factor risk premiums attributable to the risk factors in the three-factor model leads to a conclusion that the factors are jointly priced in the market. (See *Table 4.5.b.)*

Table 4.5.b Factor Risk Premiums: Joint test for the Three-Factor Model

The joint test (three factors)	5,918439677
F-value	0,000665411
The joint test (excluding the market risk factor)	0,647774187
F-value	0,524177682

[18] A description of the factor risk premiums calculation can be found in the section *3.1.5 Factor Risk Premium: Estimation Technique, Empirical Tests, and Modeling Specifics*

Table 4.5.c summarizes for the single significance testing performed on the factor risk premiums attributable to the risk factors involved in the four-factor model.

Table 4.5.c Factor Risk Premiums: Single Tests for the Four-Factor Model

	Market risk factor	Size factor	Value factor	Default factor
Factor risk premium	-0,027263106	0,010663409	0,008076965	0,007855829
Standard errors	2,57382E-05	8,94383E-05	9,08021E-05	0,000190926
Test statistic	-5,373857457	1,127545801	0,847618185	0,568538251
Critical value	3,85349E-08	0,870244106	0,801674678	0,71516523

As can be seen, the risk premiums for the size, value, and default risk factors are positive. The standard errors of estimates of the size and value factor risk premiums are very close to zero. The standard error of estimate of the default factor risk premium is somewhat higher. Only the size and value factor risk premiums are found to be significantly different from zero. Therefore, within the four-factor model, the risks underlying the size and value factors are common in equity returns and hence are considered to be priced in the market.

Testing joint significance of the factor risk premiums attributable to the risk factors in the four-factor model leads to a conclusion that the factors are jointly priced in the market. (See *Table 4.5.d.)*

Table 4.5.d Factor Risk Premiums: Joint test for the Four-Factor Model

the joint test (four factors)	7,437878005
F-value	1,21395E-05
the joint test (three factors excluding the market risk factor)	0,680470409
F-value	0,564822404

5 Analysis

This chapter analyses the empirical findings of the current study in the light of the theoretical framework and with respect to the working hypotheses.

The performed study examines the augmentation of the three-factor model of Fama and French with default risk factor. The empirical evidence of the study shows that on average the four-factor models outperform the classical three-factor models when applied on the data matherial of the NASDAQ-listed companies during the estimation period of 17 years of monthly observations[19].

In general, this conclusion is consistent with some of the prior findings, such as Vassalou's *et al* (2004). In particular, it must be noticed that the model that is defined to possess the highest econometric quality is the four-factor model that explains the average returns on the portfolio of stocks characterized by high value of market capitalization, high value of book-to-market equity, and short distance-to-default estimate[20,21]; whereas the existing literature on the three-factor model revising reports for the simple averages of dependent returns. However, the four-factor regression estimation that uses the simple averages of returns as dependent variable has not established any relationship between the candidate risk factors and dependent average returns, when performed on the data material used in the current study.

The estimation of the selected four-factor regression model that explains the average returns on the "big size – high value – short distance-to-default" portfolio provides the expected signs of the significant parameters estimates.

The negative parameter estimate for the size variable and the positive parameter estimate for the value variable are in accordance with the classical three-factor model specification of Fama *et al* (1993). The obtained positive relationship between dependent returns and default

[19] Referring to the reduced time series as argued in the subsection *4.1 The Results of an Ocular Inspection of the Time Series of Returns*
[20] The "big size – high value – short distance-to-default" portfolio
[21] The selected model

risk factor meets the expectations formulated in the working hypothesis[22] and hereby confirms the assumption that the persistently long/short distance to the default point may influence how the market reacts on changes in corporate default risk level. The firms from the working sample demonstrate long distances-to-default which implies that the default risk, these firms are exposed to, causes the expected return to rise; hence a positive sign of the parameter estimate of the default risk factor formulated as "short-minus-long" mimicking portfolio of distance-to-default.

However, if some further hypothesizing is allowed here, the same empirical finding concerning the "default risk – stock return" relationship would be hardly sensible if the firms in question demonstrated short distances-to-default over relatively long period of time. In this case, the portfolio that mimics the default risk factor would be formulated as "long-minus-short" mimicking portfolio of distance-to-default and a negative relationship "default risk – stock return" would be expected. If this was empirically proven, it would be consistent with the finding of Garlappi *et al* (2008). Their study shows that the equity returns are hump-shaped in default probability.

In this context, the firms representing the working sample of the current study are concentrated in the ascending fragment of the graph[23] depicting this hump-shaped relationship; whereas the hypothetical firms exhibiting persistently short distances-to-default would be concentrated in the descending fragment of this graph.

It must be noticed that there exist examples of mutually opposing results regarding relationship "default risk – expected return".[24] This fact confirms the specificity of default risk as regards equity returns, and suggests assessing default risk with respect to a number of conditions, such as relative distance to the default point, interplay between capital structure and firm value, overall state of economy, and not least possibility of re-negotiation upon distress.

Thus, the empirical results concerning relationship "default risk – equity return" can be seen as providing confirmation to the working hypothesis that equity returns are non-linear in default risk.

---

[22] See the subsection *3.1.3 The Working Hypothesis Regarding Relationship "Distance to Default – Average Stock Return"*

[23] Recall *Figure 1* in the subsection *3.1.3 The Working Hypothesis Regarding Relationship "Default Risk – Average Stock Return"*

[24] See the sections *1.3 Positioning of the Current Study* and *2.2 Default Risk in Equity Returns: Prior Research*

The empirical findings of the current study testify to exact pricing provided by the selected model since the estimated intercept parameter is found to be statistically zero. Further, the significance tests performed on the factor risk premiums of the size, value and default risk factors document that only the risks underlying the size and value factors are priced in the market. These risks are therefore considered to be common in equity returns and hence systematic. However, when the factor risk premiums of the corresponding size and value factors have been estimated and tested within the three-factor model in the current study, only the risk underlying the value factor is stated to be common in equity returns. For comparison, the risks underlying both the size and value factors are common in equity returns in accordance with Fama *et al* (1993).

The fact that default risk is not priced in the market and thus not systematic contradicts the finding of Vassalou *et al* (2004), but at the same time is in line with the conclusions of Gharghori *et al* (2007), disregarding the fact that these two reference studies elaborate similar empirical designs. Also, Anginer *et al* (2010) have not found any statistical support for the hypothesis that the default risk is common in equity returns. However, these authors use a marketable proxy for default risk, unlike Vassalou *et al* (2004) and Gharghori *et al* (2007) who apply structural modeling for the default risk measure obtaining.

Further, the empirical results of this study fail to provide any evidence for the size and value factors being concentrated in default risk factor. The academic discussion on the subject of economic meaning of the classical size and value factors and also criticism originated by their empirical nature, have been developing into contemporary attempts to revise the asset pricing specification of Fama and French. (Robotti, 2002) The modern researchers suspect the classical size and value factors actually being manifestations of some wider risk spectrum than the risks associated solely with relative profitability.

In this respect, the selected four-factor model cannot show evidence of concentration of size and value factors in default risk. The three-factor model expanded by default risk factor is found to be able to explain average returns on the single portfolio. At the same time, the size and value factors have not lost their explanatory power as regards equity returns after the three-factor model has been expanded. This implies that the size and value factors do not contain any default related information. On contrary, the empirical support to the idea that the size and value factors may actually contain some default related information is provided by the study by Vassalou *et al* (2004).

The findings that the size and value factors are not proxies for default related information and that the default risk is not common in equity returns, confirm the stated within this thesis working hypothesis that the Fama-French factors are expected to be proxies for systematic portion of default risk if and only if the default risk factor can be shown being priced in the market.[25]

[25] See the subsection *3.1.6 The Working Hypothesis Regarding the Size and Value Factors being Proxies for Default Related Information*

6 Conclusion

This final chapter summarizes for the conducted study and suggests some future research directions.

The empirical modeling applied to this study is premised on the three-factor asset pricing specification of Fama and French, enhanced by the default risk factor. The study utilizes the factor mimicking portfolio technique for modeling the risks underlying size, value and default risk factors. Distance-to-default estimate, deduced from the option-based model, is adopted by this study as a proxy for default risk.

The empirical results of this thesis suggest that the augmentation of the three-factor model with default risk factor improve the performance of a conventional asset pricing specification, in general. The four-factor model that aims to explain the average returns on the portfolio of stocks characterized by high value of market capitalization, high value of book-to-market equity, and short distance-to-default estimate[26] exhibits the highest explanatory properties in its class of models and in comparison to the three-factor model that aims to explain the average returns on the portfolio of stocks characterized by high value of market capitalization and high value of book-to-market equity[27].

The regression analysis of the selected model shows that the factor loadings of the portfolios of size, value and default risk factors exhibit strong properties of risk factor sensitivities for stocks. Moreover, the selected model provides exact pricing.

However, the study shows no evidence for the size and value factors being concentrated in default risk. This implies that the size and value factors are not documented being proxies for default related information.

The performed empirical tests cannot provide any evidence of systematic character of default risk. Thus, the study arrives at conclusion that default risk is not common in the

[26] The selected model
[27] The ”big size – high value” portfolio

examined equity returns and not priced in the market. At the same time, the size and value variables are found being proxies for sensitivity to common risk factors in returns.

Future work could try to analyze the possible differences in how default risk influences equity returns and asset returns. Such a study could be implemented on distressed firms in a cross-sectional setting. Further, asset volatility could be tested as a proxy for default risk within the asset pricing modeling and utilizing the mimicking portfolio technique.

References

Acton, F. S. (1970), “Numerical Methods That Work” *Washington: Mathematical Association of America*

Altman, E.I., (1968), “Financial ratios, discriminant analysis and the prediction of corporate bankruptcy” *Journal of Finance*, Vol. 23, 1968, pp. 589–609

Anginer, D. and Yıldızhan, C. (2010), “Is there a Distress Risk Anomaly? Corporate Bond Spread as a Proxy for Default Risk” (2010)

Bharath, S. T. and Shumway, T. (2004), “Forecasting Default with the KMV-Merton Model. University of Michigan” Working paper

Bharath, S. T. and Shumway, T. (2008), “Forecasting Default with the Merton Distance to Default Model” *The Review of Financial Studies*, Vol. 21, Issue 3, 2008, pp. 1339-1369

Black, F. and Scholes, M. (1973), “The Pricing of Options and Corporate Liabilities” *The Journal of Political Economy*, Vol. 81, No. 3, 1973, pp. 637-654

Bohn, J. and Crosbie, P. (2003), “Modelling Default Risk: Modelling Methodology. Moody’s KMV™” Working paper

Bonfim, D. (2009), “Credit Risk Drivers: Evaluation the Contribution of Firm Level Information and of Macroeconomic Dynamics” *Journal of Banking and Finance*, 33, 2009, pp. 281-299

Brockman, P. and Turtle, H. J. (2003), “A Barrier Option Framework for Corporate Security Valuation” *Journal of Financial Economics*, 67, 2003, pp. 511-529

Bunn, P. (2003), “Company-Accounts-Based Modelling of Business Failures and the implications for financial stability” *Bank of England*, Working paper

Campbell Y. J., Andrew W. L., and Mackinlay A.C, (1997), “The Econometrics of Financials Markets”. *Princeton University Press*, ISBN-13 978-0-691-04301-2

Chan-Lau, Jorge A. and Yinqiu Lu, (2006), “Idiosyncratic and Systematic Risk in the European Corporate Sector: A CDO Perspective,” *IMF Working Paper,* 06/107, (Washington: International Monetary Fund).

Chen, N.-F., Roll, R., Ross, S. A., (1986), “Economic Forces and the Stock Market” *Journal of Business,* Vol. 59, (July 1986), pp. 383–403

Chung, Y. P., Johnson, H., Schill, M. J., (2006), “Asset Pricing when Returns are Nonnormal: Fama-French Factors versus Higher-Order Systematic Comoments” *Journal of Business,* Vol. 79, pp. 923 – 940

Da, Z. and Kellogg, P. G., (2006), “Clientele Change, Liquidity Shock, and the Return on Financially Distressed Stocks” *School of Management Northwestern University*

Demchuk, A., Gibson, R., (2005), ”Stock Market Performance and the Term Structure of Credit Spreads” *University of Zürich, Swiss Banking Institute*

Dichev, I. D., (1998), “Is the Risk of Bankruptcy a Systematic Risk?” *The Journal of Finance*, Vol. 53, No. 3, (Jun., 1998), pp. 1131-1147

Dionne, G., Laajimi, S., Mejri, S. and Petrescu, M., (2008), “Estimation of the Default Risk of Publicly Traded Companies: Evidence from Canadian Data” *Canadian Journal of Administrative Science*, 25, (2008), pp. 134-152

Fama, E. F., French, K. R., (1992), “The Cross-Section of Expected Stock Returns” *The Journal of Finance,* Vol. 47, (June 1992), pp. 427–465

Fama, E. F., French, K. R., (1993), “Common Risk Factors in the Returns on stocks and Bonds” *The Journal of Financial Economics*, Vol. 33, (1993), pp. 3–56

Fama, E. F., French, K. R., (1995), “Size and Book-to-Market Factors in Earnings and Returns” *Journal of Finance,*Vol. 50, pp. 131 – 150

Ferguson, M., Shockley, R., (2003), “Equilibrium Anomalies” *Journal of Finance,* Vol. 58, pp. 2549 – 2580

Ferson, W. E., Campbell R. H. (1991), “The Variation of Economic Risk Premiums” *Journal of Political Economy,* Vol. 99 (April 1991), pp. 385–415

Garlappi, L. and Hong Y., (2008), ”Financial Distress and the Cross Section of Equity Returns” *Review of Financial Studies forthcoming*.

Gertner, R. and Scharfstein, D., (1991), “A Theory of Workouts and the Effects of Reorganization Law” *The Journal of Finance*, Vol. 46, No. 4, (1991), pp. 1189-1222

Gersbach, H. and Lipponer, A., (2003), “Firm Defaults and Correlation Effect” *European Financial Management,* Vol. 9, No. 3, (2003), pp. 361-377

Gharghori P., Howard C. and Robert F., (2009), “Default risk and equity returns: Australian evidence” *Pacific-Basin Finance Journal*, Vol. 17, (2009), pp. 580–593

Harris, R., Sollis, R., (2003), “Applied Time Series Modelling and Forecasting” *Wiley*

Huberman, G., Kandel, S., Stambaugh, R. F., (1987), "Mimicking Portfolios and Exact Arbitrage Pricing” *Journal of Finance,* Vol. 42(1), pp. 1 – 9

Kaw, A. (2009), “Text book notes of Newton-Raphson method of finding roots of nonlinear equation, including convergence and pitfalls” (http://numericalmethods.eng.usf.edu)

Korablev, I. and Jiang, Y. (2007), “Default Case Study” *Moody’s KMV,* (April 2007)

Merton, R. C. (1973), “An Intertemporal Capital Asset Pricing Model” *Econometrica,* Vol. 41, (September 1973), pp. 867–87

Merton, R. C., (1974), “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates” *The Journal of Finance*, Vol. 29, No. 2, pp. 449-470

Nesterov, Y., (2004), “Introductory Lectures on Convex Optimization. A Basic Course” *Springer Netherlands*, ISBN10: 1402075537

Ogden, J.P., Jen, F.C. and O’Connor, P.F., (2003), “Advanced Corporate Finance: Policies and Strategies” *Prentice Hall*

Pennacchi, G., (2008), “Theory of Asset Pricing” *Pearson*

Pesaran, M. H. and Schuermann, T., (2003), “Credit Risk and Macroeconomic Dynamics” *University of Cambridge*, *Federal Reserve Bank of New York*

Pettit, J., (2007), “Strategic Corporate Finance: Applications in Valuation and Capital Structure” *Wiley.*

Qu, Y., (2008), “Macro Economic Factors and Probability of Default” *European Journal of Econometrics*, Finance and Administrative Science, Vol.13

Rietz, T. A., (1988), “The Equity Risk Premium, A Solution” *Journal of Monetary Economics,* Vol. 22, pp.117–131

Ross, S.A., (1976), “The arbitrage theory of capital asset pricing” *Journal of Economic Theory*, Vol. 13, (1976), pp. 341–60.

Saunders, A. and Allen, L., (2002), “Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms” John Wiley & Sons, Inc.

Sharpe, W., (1963) “A Simplified model for Portfolio Analysis“ *The Management Science,* (Jan 1963), Vol. 9, Issue 2, (1963), pp. 277-293

Sharpe, W., (1964), “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, Vol. 19, (1964), pp. 425–42.

Sivokobylenko, V. F., (2007), “Mathematical Problems of Electrical Engineering” (Сивокобыленко, В. Ф., (2007), “Математические задачи электротехники”) *Donetsk National Technical University*

Vassalou, M. and Yuhang X., (2004), “Default Risk in Equity Returns” *Journal of Finance*, Vol. 59, (2004), pp. 831–68.

Wruck, K. H. (1990), “Financial Distress, Reorganization, and Organizational Efficiency” *Journal of Financial Economics*, Vol. 27, (1990), pp. 419-444

Appendix I

I.1 The Time Series of the Simple Averages of the Distance-to-Default Estimates of all Firms: Descriptive Statistics

Table I.1. Descriptive Statistics for the Time Series Observations on The Simple Averages of the Distances-to-Default of all Firms

	DISTANCE-TO-DEFAULT
Mean	829,4213
Median	822,3717
Maximum	1858,632
Minimum	427,5763
Std. Dev.	326,8564
Skewness	1,408929
Kurtosis	5,883238
Jarque-Bera	13,54448
Probability	0,001145
Observations	20

Appendix II

II.1 The Gradient and Newton-Raphson Methods for Solving the Specified System of Non-Linear Equations

In order to solve the system of non-linear equations simultaneously, the programming capabilities of MathCAD have been employed. The two different methods have been used – the gradient method (Nesterov, 2004; Sivokobylenko, 2007) and the Newton-Raphson method (Sivokobylenko, 2007; Kaw, 2009), in order to control for the possible errors implied by the potential shortcomings of each method.

Thus, not going too deeply into the technical interpretation of the methods, it may be noticed that they exploit the same idea of dynamic optimization. The first method is built on the narrowing of the initial value to the sought root, where every iteration step is defined from the differentiation of the Taylor raw of the non-linear components with respect to the gradient of the function. The second method can be explained geometrically and exploits the procedure of adjustments to the initially stated value by finding the optimal intersection of the function's tangent and the axis of abscissae. (Acton, 1970; Nesterov, 2004; Sivokobylenko, 2007)

However, the two methods exhibit different kinds of weaknesses. The gradient method guarantees the convergence of calculations, but sometimes it may imply difficulties regarding the iteration parameter which must be defined for every step. (Nesterov, 2004; Sivokobylenko, 2007) The Newton-Raphson method, on the other hand, may lead to the root-jumps, despite the accuracy of initial values stated; however, the problem associated with the iteration parameter is mitigated here since this method controls for oscillation in function. (Sivokobylenko, 2007; Kaw, 2009)

The calculation has shown that the two methods employed, generate the same results, differing only sometimes and insignificantly. This can be probably explained, beyond just stating that the calculation results are reliable, by the fact that the equations of interest are "almost linearly" related, since the obtained values for $N(d_1)$ and $N(d_2)$ are units.

II.2 The Input Data Modifications when Extracting Asset Value and Asset Volatility from Market Value of Equity and Equity Volatility

It is worth noticing that some modifications of the input data have been performed before solving the system of equations. Since it is not desirable to have any output data falling off, all zeros for the default point have been replaced by 1E-10. Algebraically, this is judged to be fairly unacceptable since the default points of zero cannot be used in the calculations (since appearing in denominator of the formulas), and the mentioned replacement leads to inadequate results, namely to infinitely large values.

However, appealing to the financial theory and with respect to the further use of the output data, this replacement is rather appropriate. The structural interpretation of default risk is based on the capital structure, which defines the value of assets and asset volatility and hence individual distances-to-default. In this setting, the default point of zero (which in the option vocabulary is a strike price) implies that equity value equals assets value, and that there is no threat of defaulting. The mentioned replacement also leads to the absence of threat of default since the model in this case generates the distance-to-default estimates which are infinitely large (infinitely long distances-to-default).

II.3 On Solving the Specified System of Non-Linear Equations in MathCAD

MathCAD is computer software primarily used for engineering calculations. Its utilization in this thesis assignment allows for solving of the system of non-linear equations without any technical complications. The built-in programming capabilities of MathCAD are rather intuitive and visual. However, the algebraic understanding of solving process is a requirement since a calculation procedure to some extend depends on how the problem is formulated for the software program. Below one can find a visual illustration of performed calculations supported by a short explanation of the necessary commands.

The solving procedure starts with introduction of the variables expressed in vector form. A sign of equality placed after the introduced variable is obtained by using the following symbol of punctuation “:” (colon). Thus, the variables involved in the equation system and the vectors of corresponding variables are introduced and shown here:

$X =$

	0
0	$1.208\cdot10^{3}$
1	712.5
2	$6.606\cdot10^{3}$
3	$1.326\cdot10^{3}$
4	$1.011\cdot10^{3}$
5	$5.52\cdot10^{4}$
6	$5.5\cdot10^{4}$
7	$5.5\cdot10^{4}$
8	$2.995\cdot10^{4}$
9	$2.977\cdot10^{4}$
10	$2.183\cdot10^{4}$
11	$1.521\cdot10^{3}$
12	$1.356\cdot10^{5}$
13	346
14	$1\cdot10^{-10}$
15	$1\cdot10^{-10}$

$V_E =$

	0
0	$2.328\cdot10^{5}$
1	$3.06\cdot10^{5}$
2	$8.176\cdot10^{5}$
3	$1.45\cdot10^{6}$
4	$3.354\cdot10^{6}$
5	$6.456\cdot10^{6}$
6	$1.239\cdot10^{7}$
7	$6.232\cdot10^{6}$
8	$1.608\cdot10^{7}$
9	$1.681\cdot10^{7}$
10	$3.004\cdot10^{6}$
11	$2.197\cdot10^{6}$
12	$1.656\cdot10^{6}$
13	$3.005\cdot10^{6}$
14	$1.638\cdot10^{6}$
15	$1.387\cdot10^{6}$

$\sigma_E =$

	0
0	0.033
1	0.013
2	0.037
3	0.04
4	0.087
5	0.106
6	0.114
7	0.127
8	0.057
9	0.083
10	0.264
11	0.022
12	$7.576\cdot10^{-3}$
13	0.012
14	0.014
15	$2.793\cdot10^{-3}$

$r_f =$

	0
0	0.074
1	0.055
2	0.037
3	0.033
4	0.05
5	0.056
6	0.052
7	0.053
8	0.048
9	0.048
10	0.058
11	0.038
12	0.038
13	0.032
14	0.032
15	0.029

$T = 1$

Hereafter, the range variable k must be defined. It is (in the simple cases) dependent on the length of the input vectors. In the current example expression *k:=0..579* should be read as “the calculations will be performed in 579 steps”. The sign “..” is obtained by pressing semicolon-button on the tangent board.

Next, the initial values are stated ($\sigma_A = 0.5\ and\ V_A = 1000000$), and the system of equations is introduced by command “Given”. When writing the system of equations, it must be kept in mind that the equality signs used in equation specifications are so called “symbolic equality signs”, which means that the equality is not obvious, but must be achieved in the solving process. This “symbolic equality sign” is obtained by pressing Ctrl and “=” simultaneously.

In order to invoke any function, such as “Given”, “pnorm” or “Minerr”, the Insert-button on the program toolbar must be used and the *f(x)*-symbol must be chosen; thereafter it remains to define which function is to be applied.

$k := 0..579$

$\sigma_A := 0.5$ $V_A := 100000$

Given

$$V_A \cdot \text{pnorm}\left(\frac{\ln\left(\frac{V_A}{X_k}\right) + r_{f_k} + \frac{1}{2}\cdot\sigma_A{}^2\cdot T}{\sigma_A\cdot\sqrt{T}}, 2, 1\right) - e^{-r_{f_k}\cdot T}\cdot X_k\cdot\text{pnorm}\left[\frac{\ln\left(\frac{V_A}{X_k}\right) + \left[r_{f_k} + \frac{1}{2}\cdot(\sigma_A)^2\right]\cdot T}{\sigma_A\cdot\sqrt{T}} - \sigma_A\cdot\sqrt{T}, 2, 1\right] - V_{E_k} = 0$$

$$\frac{V_A}{V_{E_k}}\cdot\text{pnorm}\left[\frac{\ln\left(\frac{V_A}{X_k}\right) + \left[r_{f_k} + \frac{1}{2}\cdot(\sigma_A)^2\right]\cdot T}{\sigma_A\cdot\sqrt{T}}, 2, 1\right]\cdot\sigma_A - \sigma_{E_k} = 0$$

It must be noticed that the theoretical equations are transformed such that the minimization is presupposed: the distances between actual V_E and its algebraic function, and between actual σ_E and its algebraic function are aimed to be minimized in order to guarantee optimization.

Now, it is possible to define the matrix that will contain the output. The program can only recognize the Z-matrix with k-coefficient. The sought vectors of variables must be introduced by using function "Minerr". Thus, the expression $Z_k \coloneqq Minerr(V_A, \sigma_A)$ should be read as "the matrix of output Z_k contains the columns of sought variables V_A and σ_A". Also, the clarification on how the program should seek the variables V_A and σ_A is necessary. For this purpose, it is enough to state in how many steps the equation system must be solved by writing $(V_A\_0\_580)_k \coloneqq (Z_k)_0$ and $(\sigma_A\_0\_580)_k \coloneqq (Z_k)_1$. Below follows a corresponding clip from the MathCAD working sheet.

$Z_k := \text{Minerr}(V_A, \sigma_A)$

$(Va\_0\_580)_k := (Z_k)_0$ $(\sigma a\_0\_580)_k := (Z_k)_1$

In order to solve the system of equations, the sought variables must be written once more followed by the "true equality signs" (obtained by pressing colon-button on the tangentboard).

$Va\_0\_580_k =$

	0
0	$2.339 \cdot 10^{5}$
1	$3.067 \cdot 10^{5}$
2	$8.24 \cdot 10^{5}$
3	$1.451 \cdot 10^{6}$
4	$3.355 \cdot 10^{6}$
5	$6.508 \cdot 10^{6}$
6	$1.244 \cdot 10^{7}$
7	$6.284 \cdot 10^{6}$
8	$1.611 \cdot 10^{7}$
9	$1.684 \cdot 10^{7}$
10	$3.025 \cdot 10^{6}$
11	$2.198 \cdot 10^{6}$
12	$1.787 \cdot 10^{6}$
13	$3.005 \cdot 10^{6}$
14	$1.638 \cdot 10^{6}$
15	$1.387 \cdot 10^{6}$

$\sigma a\_0\_580_k =$

	0
0	0.033
1	0.013
2	0.037
3	0.04
4	0.087
5	0.105
6	0.114
7	0.126
8	0.057
9	0.083
10	0.262
11	0.022
12	$7.022 \cdot 10^{-3}$
13	0.012
14	0.014
15	$2.793 \cdot 10^{-3}$

The optimization method for solving the system of equations can be chosen by a right click on the function “Minerr”. For the equation system in question both *Conjugate Gradient* method and *Quasi-Newton* method are appropriate.

II.4 Some Details on Dealing with the Output Data as Regards Distance-to-Default Estimates

Recalling the discussion concerning the default points of zero, the two ways of dealing with output data for the distances-to-default of the corresponding firms are possible. It might be recommended to exclude these infinitely large output values for distances-to-default from the sample, if one is interested in actual distances-to-default. In this case the corresponding firms must be seen as those which are not possible to be defaulting in the structural setting.

On the other hand, the obtained infinitely large values of distances-to-default for the firms which default points are zero, may not be excluded from the output data, if one is not interested in actual distances-to-default, but rather in the variation in the distance-to-default-variable. The second option is suitable for this thesis assignment. That is because the time-series of distances-to-default are not explicitly used in this study. Instead, it is just variation in the variable that matters, since the values of the distance-to-default-variable are used for portfolio formation and are the criterions for sorting of the stocks into different portfolios. Thus, the infinitely large distances-to-default end up in the portfolios of the highest sample percentile.

Also, as distance-to-default estimate is simply a standard deviation of the asset value from the default point, or expressed algebraically, a distance to the origin (if translating an assumed normal distribution to a standard normal distribution), the obtained negations of distance-to-default estimates must be replaced by the corresponding absolute values for the further calculations.

Appendix III

III.1 The Graphs of the Time Series

Figure III.1. The graphs of the Time Series of Explanatory and Dependent Returns – Entire Sample

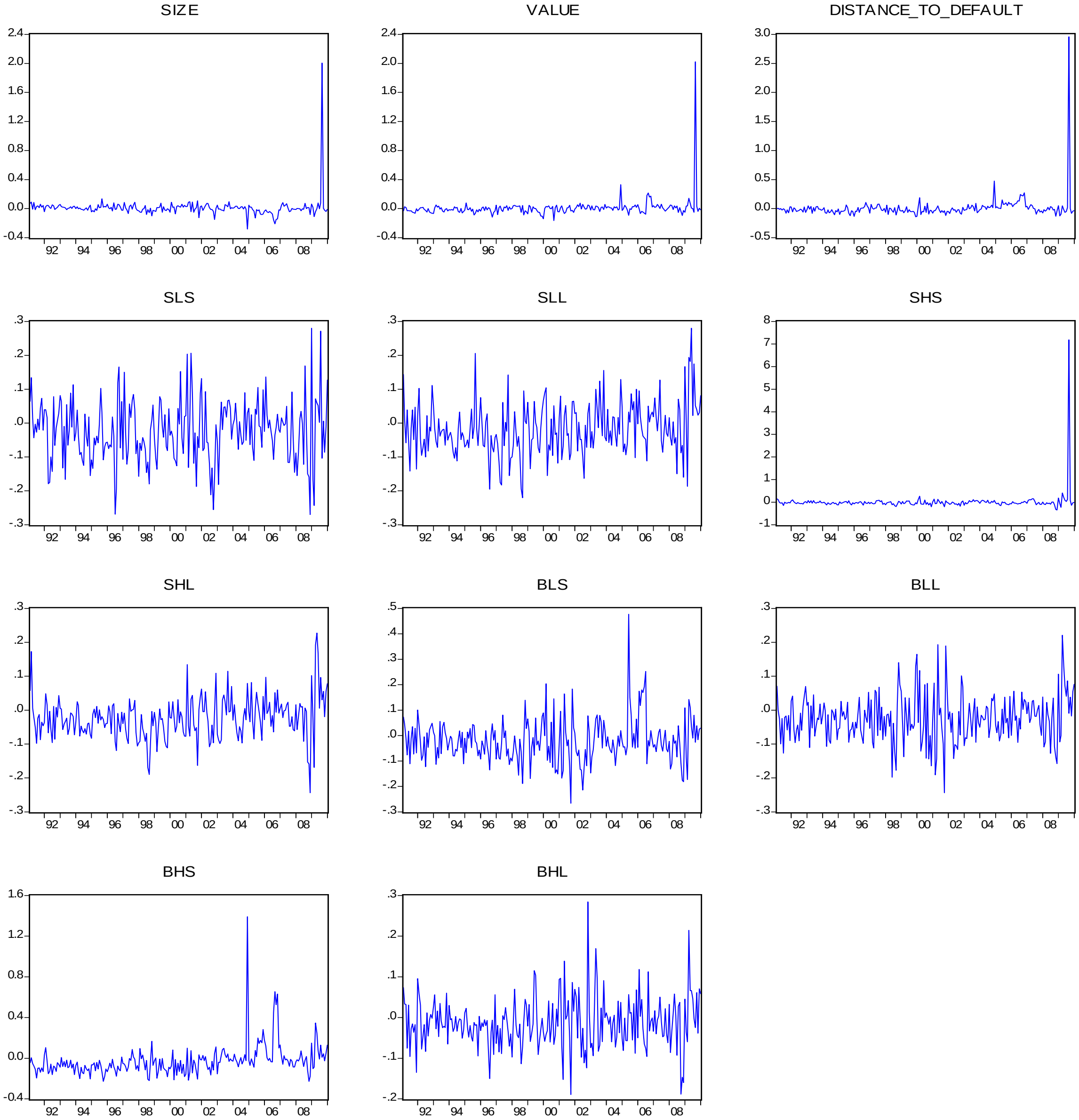


Figure III.1 illustrates the graphs of the time series of the explanatory variables Size, Value and Distance to Default alongside with the time series of the dependent variables formed on the returns of the portfolios representing the firms with certain characteristics: “small size – low value – short distance-to-default” (SLS), “small size – low value – long distance-to-default” (SLL), “small size – high value – short distance-to-default” (SHS), “small size – high value – long distance-to-default” (SHL), “big size – low value – short distance-to-default” (BLS), “big size – low value – long distance-to-default” (BLL), “big size – high value – short distance-to-default” (BHS), “big size – high value – long distance-to-default” (BHL). The length of the time series lies between February, 1991 and January, 2010.

Figure III.2. The graphs of the Time Series of Explanatory and Dependent Returns – Sub-Sample

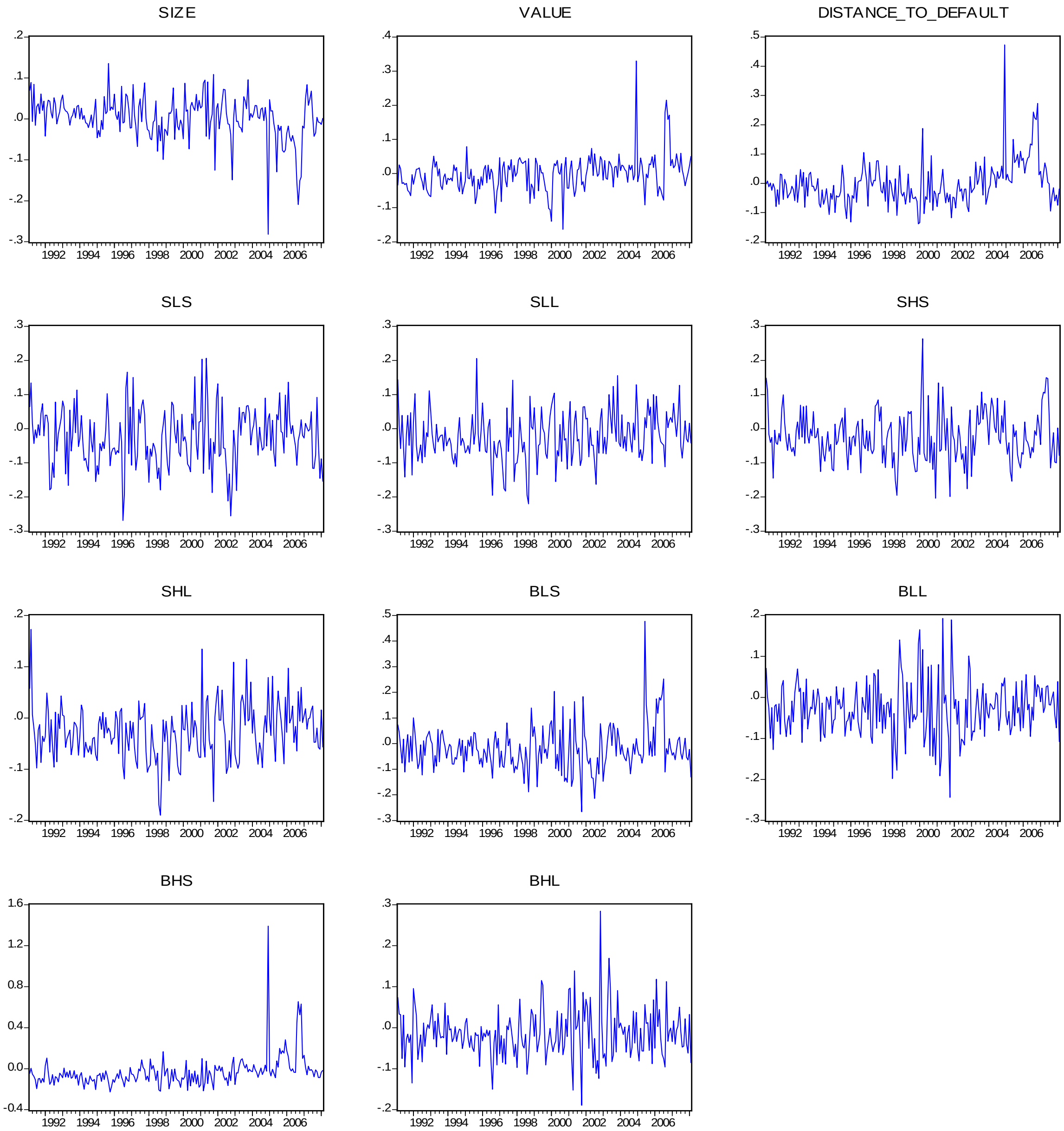


Figure III.2 illustrates the graphs of the time series of the explanatory variables Size, Value and Distance to Default alongside with the time series of the dependent variables formed on the returns of the portfolios representing the firms with certain characteristics: "small size – low value – short distance-to-default" (SLS), "small size – low value – long distance-to-default" (SLL), "small size – high value – short distance-to-default" (SHS), "small size – high value – long distance-to-default" (SHL), "big size – low value – short distance-to-default" (BLS), "big size – low value – long distance-to-default" (BLL), "big size – high value – short distance-to-default" (BHS), "big size – high value – long distance-to-default" (BHL). The length of the time series is from February 1991 to February 2008.

III.2 Descriptive Statistics of the Time Series

Table III.1. Descriptive Statistics for the Time Series Involved in the Three-Factor Model Estimation – Entire Sample

	RM	SIZE	VALUE	SL	SH	BL	BH
Mean	-0,03	0,00	0,01	-0,02	-0,01	-0,02	-0,02
Median	-0,02	0,00	0,00	-0,02	-0,03	-0,02	-0,04
Maximum	0,18	1,10	1,65	0,21	2,42	0,22	0,88
Minimum	-0,26	-0,47	-0,20	-0,21	-0,29	-0,26	-0,20
Std. Dev.	0,07	0,09	0,14	0,07	0,17	0,07	0,11
Skewness	-0,29	6,37	8,12	0,16	11,91	0,26	3,41
Kurtosis	4,06	86,10	91,77	3,41	166,57	3,74	24,99
Jarque-Bera	13,86	67148,82	77359,20	2,59	259578,80	7,79	5035,83
Probability	0,00	0,00	0,00	0,27	0,00	0,02	0,00

Table III.1 summarizes for the Descriptive Statistics for the time series of explanatory variables Market risk factor, Size factor and Value factor alongside with the time series of dependent variables formed on the returns of the portfolios representing the firms with the following characteristics: “small size – low value” (SL), “small size – high value” (SH), “big size – low value” (BL), “big size – high value” (BH). The descriptive statistic tests are applied on the entire sample (from February 1991 to January 2010). The table contains means, maximum values, the minimum values, standard deviations, skewness, kurtosis and the Jarque-Bera test statistics.

Table III.2. Descriptive Statistics for the Time Series Involved in the Three-Factor Model Estimation – Sub-Sample

	RM	SIZE	VALUE	SL	SH	BL	BH
Mean	-0,03	0,00	0,00	-0,02	-0,03	-0,02	-0,03
Median	-0,02	0,00	0,00	-0,02	-0,03	-0,03	-0,04
Maximum	0,18	0,16	0,76	0,15	0,15	0,22	0,88
Minimum	-0,26	-0,47	-0,20	-0,21	-0,19	-0,26	-0,20
Std. Dev.	0,07	0,06	0,09	0,07	0,05	0,07	0,11
Skewness	-0,37	-2,72	3,72	-0,03	0,14	0,30	3,82
Kurtosis	4,39	21,09	32,28	2,95	3,56	4,03	28,51
Jarque-Bera	21,22	3047,47	7794,98	0,05	3,32	12,06	6056,81
Probability	0,00	0,00	0,00	0,97	0,19	0,00	0,00

Table III.2 summarizes for the Descriptive Statistics for the time series of explanatory variables Market risk factor, Size factor and Value factor alongside with the time series of dependent variables formed on the returns of the portfolios representing the firms with the following characteristics: “small size – low value” (SL), “small size – high value” (SH), “big size – low value” (BL), “big size – high value” (BH). The descriptive statistic tests are applied on the sub-sampled data (from February 1991 to February 2008). The table contains means, maximum values, the minimum values, standard deviations, skewness, kurtosis and the Jarque-Bera test statistics.

Table III.3. Descriptive Statistics for the Time Series Involved in the Four-Factor Model Estimation – Entire Sample

	RM	SIZE	VALUE	DEFAULT	SLS	SLL	SHS	SHL	BLS	BLL	BHS	BHL
Mean	-0,03	0,01	0,01	0,01	-0,02	-0,02	0,01	-0,02	-0,02	-0,02	-0,03	-0,01
Median	-0,02	0,00	0,00	-0,02	-0,03	-0,03	-0,03	-0,02	-0,03	-0,02	-0,04	-0,02
Maximum	0,18	2,00	2,02	2,96	0,28	0,28	7,18	0,23	0,48	0,22	1,39	0,28
Minimum	-0,26	-0,28	-0,16	-0,14	-0,27	-0,22	-0,35	-0,24	-0,27	-0,24	-0,23	-0,19
Std, Dev,	0,07	0,14	0,14	0,21	0,09	0,08	0,48	0,06	0,09	0,07	0,16	0,06
Skewness	-0,29	11,96	12,13	12,47	0,17	0,45	14,30	0,43	1,08	0,31	4,48	0,61
Kurtosis	4,06	168,84	169,56	175,66	3,61	3,95	212,14	5,24	7,51	3,98	35,71	5,49
Jarque-Bera	13,86	266721	269148	289129	4,66	16,34	423310	54,49	237,54	12,82	10923	72,93
Probability	0,00	0,00	0,00	0,00	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Table III.3 summarizes for the Descriptive Statistics for the time series of explanatory variables Market risk factor, Size factor, Value factor, and Default risk factor alongside with the time series of dependent variables formed on the returns of the portfolios representing the firms with the following characteristics: "small size – low value – short distance-to-default" (SLS), "small size – low value – long distance-to-default" (SLL), "small size – high value – short distance-to-default" (SHS), "small size – high value – long distance-to-default" (SHL), "big size – low value – short distance-to-default" (BLS), "big size – low value – long distance-to-default" (BLL), "big size – high value – short distance-to-default" (BHS), "big size – high value – long distance-to-default" (BHL). The descriptive statistic tests are applied on the entire sample (from February 1991 to January 20010). The table contains means, maximum values, the minimum values, standard deviations, skewness, kurtosis and the Jarque-Bera test statistics.

Table III.4. Descriptive Statistics for the Time Series Involved in the Four-Factor Model Estimation – Sub-Sample

	RM	SIZE	VALUE	DEFAULT	SLS	SLL	SHS	SHL	BLS	BLL	BHS	BHL
Mean	-0,03	0,00	0,00	0,00	-0,03	-0,02	-0,02	-0,03	-0,02	-0,03	-0,03	-0,01
Median	-0,02	0,01	0,00	-0,01	-0,03	-0,03	-0,03	-0,03	-0,03	-0,03	-0,05	-0,02
Maximum	0,18	0,14	0,33	0,47	0,21	0,21	0,26	0,17	0,48	0,19	1,39	0,28
Minimum	-0,26	-0,28	-0,16	-0,14	-0,27	-0,22	-0,20	-0,19	-0,27	-0,24	-0,23	-0,19
Std. Dev.	0,07	0,05	0,05	0,07	0,08	0,07	0,07	0,05	0,09	0,07	0,16	0,06
Skewness	-0,37	-1,34	1,54	2,04	0,01	0,18	0,41	0,32	1,25	0,21	4,78	0,80
Kurtosis	4,39	7,77	11,72	11,97	3,09	3,40	3,73	4,18	8,17	4,11	38,01	5,92
Jarque-Bera	21,22	256,00	729,85	829,56	0,08	2,52	10,20	15,29	281,69	11,98	11250	94,54
Probability	0,00	0,00	0,00	0,00	0,96	0,28	0,01	0,00	0,00	0,00	0,00	0,00

Table III.4 summarizes for the Descriptive Statistics for the time series of explanatory variables Market risk factor, Size factor, Value factor, and Default risk factor alongside with the time series of dependent variables formed on the returns of the portfolios representing the firms with the following characteristics: "small size – low value – short distance-to-default" (SLS), "small size – low value – long distance-to-default" (SLL), "small size – high value – short distance-to-default" (SHS), "small size – high value – long distance-to-default" (SHL), "big size – low value – short distance-to-default" (BLS), "big size – low value – long distance-to-default" (BLL), "big size – high value – short distance-to-default" (BHS), "big size – high value – long distance-to-default" (BHL). The descriptive statistic tests are applied on the sub-sampled data (from February 1991 to February 2008). The table contains means, maximum values, the minimum values, standard deviations, skewness, kurtosis and the Jarque-Bera test statistics.

III.3 Results of the Regression Estimations (the Entire Sample)

Table III.5. The Three-Factor Model Estimation Results – Entire Sample

	SL	SH	BL	BH
Alpha	**0,000**	**0,002**	**0,002**	**-0,004**
Std. Error	0,004	0,002	0,003	0,004
t-statistic	-0,137	0,878	0,842	-1,010
R_m	**0,725**	**0,899**	**0,881**	**0,788**
Std. Error	0,047	0,031	0,033	0,056
t-statistic	15,428	28,558	26,630	14,156
Size	**0,213**	**1,050**	**-0,039**	**-0,629**
Std. Error	0,039	0,026	0,027	0,046
t-statistic	5,509	40,528	-1,422	-13,720
Value	**-0,084**	**0,732**	**-0,008**	**0,474**
Std. Error	0,026	0,017	0,018	0,031
t-statistic	-3,208	41,833	-0,425	15,317
R^2	**0,537**	**0,963**	**0,766**	**0,702**

Table III.5 reports the estimation results of the Three-factor model, where the returns on portfolios “small size – low value (SL), “small size – high value” (SH), “big size – low value” (BL), “big size – high value” (BH) are regressed on the Market risk factor, Size factor and Value factor. The estimation period lies between February 1991 and January 2010. The table presents the estimated coefficients of intercepts and parameters for each regression alongside with the standard errors, t – statistics and the values of R–squared.

Table III.6. The Four-Factor Model Estimation Results – Entire Sample

	SLL	SLH	SHL	SHH	BLL	BLH	BHL	BHH
Alpha	**-0,006**	**0,002**	**0,006**	**-0,008**	**0,003**	**-0,001**	**-0,003**	**0,002**
Std. Error	0,005	0,004	0,004	0,003	0,004	0,002	**0,006**	0,003
t-statistic	-1,382	0,590	1,490	-2,413	0,933	-0,375	**-0,539**	0,737
R_m	**0,639**	**0,770**	**1,023**	**0,607**	**0,770**	**0,862**	**0,671**	**0,598**
Std. Error	0,066	0,056	0,056	0,046	0,049	0,032	**0,078**	0,047
t-statistic	9,751	13,681	18,205	13,070	15,780	26,727	**8,622**	12,800
Size	**0,234**	**0,210**	**1,636**	**0,076**	**-0,113**	**-0,020**	**-1,394**	**-0,027**
Std. Error	0,057	0,049	0,049	0,041	0,043	0,028	**0,068**	0,041
t-statistic	4,067	4,241	33,188	1,859	-2,638	-0,711	**-20,414**	-0,668
Value	**-0,651**	**0,030**	**1,221**	**0,210**	**-0,490**	**0,017**	**0,648**	**0,355**
Std. Error	0,082	0,071	0,071	0,058	0,061	0,041	**0,098**	0,059
t-statistic	-7,902	0,425	17,271	3,592	-7,984	0,422	**6,618**	6,051
Default	**0,287**	**-0,125**	**0,458**	**-0,155**	**0,398**	**0,001**	**0,556**	**-0,228**
Std. Error	0,062	0,053	0,053	0,044	0,046	0,031	**0,074**	0,044
t-statistic	4,626	-2,346	8,597	-3,531	8,609	0,032	**7,533**	-5,148
R^2	**0,482**	**0,473**	**0,987**	**0,445**	**0,684**	**0,787**	**0,748**	**0,450**

Table III.6 reports the estimation results of the Four-factor model, where the returns on portfolios “small size – low value – short distance-to-default” (SLS), “small size – low value – long distance-to-default” (SLL), “small size – high value – short distance-to-default” (SHS), “small size – high value – long distance-to-default” (SHL), “big size – low value – short distance-to-default” (BLS), “big size – low value – long distance-to-default” (BLL), “big size – high value – short distance-to-default” (BHS), “big size – high value – long distance-to-default” (BHL) are regressed on the Market risk factor, Size factor, Value factor, and Default risk factor. The estimation period lies between February 1991 and January 2010. The table presents the estimated coefficients of intercepts and parameters for each regression alongside with the standard errors, t –statistics and the values of R–squared.