COPENHAGEN BUSINESS SCHOOL

Asset Pricing in Germany: An Empirical Study of Factor Models

by

Christoffer Iversen

Student number: 27314

Supervisor: Domenico Tripodi

Date of submission: 14 May 2018
Number of characters incl. spaces: 164012
Number of pages: 80

A thesis submitted in partial fulfillment for the
degree of MSc in Economics and Business Administration

in

Finance and Investments

May 2018

COPENHAGEN BUSINESS SCHOOL

Abstract

by Christoffer Iversen

This thesis uses a sample of German stocks traded on German stock exchanges to study the cross-section of average stock returns. Six different factors are constructed from scratch and proposed as potential explanatory variables able to explain the variation in average returns. These factors are related to different firm characteristics that potentially pose risk. In particular, size, value, momentum, profitability, investment, and covariation with the market portfolio are used in different combinations in the search for a factor model that is able to adequately explain return variations. The Fama and French [1993] three-factor model, Carhart [1997] four-factor model, Fama and French [2015] five-factor model, and a six-factor model that combines all factors are analyzed and tested using Fama and MacBeth [1973] regressions and the Gibbons, Ross, and Shanken [1989] test. This thesis finds that the five- and six-factor models are superior to the three- and four-factor models. The models are able to explain variations in average returns, but they seem to have one minor weakness related to firms with low book-to-market ratios. Five out of six risk premia are not strong enough, and a more parsimonious model may be a one-factor model which only includes the market risk premium.

Acknowledgements

I would like to express my appreciation and gratitude towards my supervisor, Domenico Tripodi, who has been a great sparring partner with constructive, rewarding, and worthwile discussions. Your advice has been very useful and a great help.

I would also like to thank my mother and father for keeping me motivated through tough times. A special thanks to my father who has an engineering background, and therefore he has been able to give me feedback from a different perspective.

Lastly, I would like to thank my fellow students at the Finance and Investments program at Copenhagen Business School who have been in a similar situation and willing to discuss various thesis related issues.

Contents

List of Figures

List of Tables

Abbreviations

APT	**A**rbitrage **P**ricing **T**heory
AUD	**A**ustralian **D**ollar
CAPM	**C**apital **A**sset **P**ricing **M**odel
CMA	**C**onservative **M**inus **A**ggressive
CRSP	**C**enter for **R**esearch in **S**ecurity **P**rices
DEM	**D**eutsche **M**ark
e.g.	**e**xempli **g**ratia (for example)
EMR	**E**xcess **M**arket **R**eturn
ETF	**E**xchange-**T**raded **F**und
EUR	**Eur**o
FRED	**F**ederal **R**eserve **E**conomic **D**ata
HML	**H**igh **M**inus **L**ow
i.e.	**i**d **e**st (that is)
GBP	**G**reat **B**ritain **P**ound
GMM	**G**eneralized **M**ethod of **M**oments
GRS	**G**ibbons **R**oss **S**hanken
NBER	**N**ational **B**ureau of **E**conomic **R**esearch
RMW	**R**obust **M**inus **W**eak
SEC	**S**ecurities and **E**xchange **C**ommission
SIC	**S**tandard **I**ndustry **C**lassification
SMB	**S**mall **M**inus **B**ig
S&P	**S**tandard and **P**oor
US	**U**nited **S**tates
USD	**U**nited **S**tates **D**ollar
WML	**W**inners **M**inus **L**osers

Chapter 1

Introduction

1.1 Background

In finance, it has long been understood that a constant stream of information arrives every day which influences supply and demand for financial securities and may contribute to variations in security prices and associated returns on a day to day basis. But when returns are assessed on an average basis, why do these vary across different securities? what determines the prices of financial assets? can we predict returns? Asset pricing models may contribute to clarifying these question. One essential piece of knowledge from asset pricing is that investors require return from bearing systematic risk. What constitutes systematic risk is another, challenging question. How do we measure the risks of financial assets? One good thing about finance is that it is an area of economics which is rich in naturally generated data. The fact that data is naturally generated can also be a challenge because researchers cannot control how much data is generated or the random fluctuations and shocks that influence the data. Empirical studies and analyses are therefore influential in the development of new insights and important theories in asset pricing. Empirical studies have contributed to the findings of some of the most well-known puzzles or paradoxes in finance.

Theorists develop models with testable predictions; empirical researchers document "puzzles" - stylized facts that fail to fit established theories - and this stimulates the development of new theories. [Campbell, 2000].

This quote captures the point of why empirical finance is interesting and different from testable theory quite well. In recent years, certain systematic risk factors have displayed great interest in the empirical finance community. This thesis is a study that put some of these factors into perspective.

The prices and returns of many assets can be observed. Empirical finance is the art of understanding why prices and returns are the way they are, and asset pricing theory can help us understand that. If the empiri is not in accordance with a theoretical model, one can say that the model is wrong and attempt to improve the model, but there is also another way of viewing things. One can instead say that the empiri is wrong. This would mean that certain assets do not have the price they should have, and that investors could theoretically exploit these mis-pricings for abnormal gains. These two ways of viewing the universe are often debated and commonly known as the positive and normative views. The normative use of theoretical models is popularly applied in practice.

Generally, the value of an investment or a financial asset depends on time and risk. The time delay from the point of buying to the point of payoff or selling is of course something that an investor should be compensated for, but much more interesting is the risk of the asset's payoff. The uncertainty of future payoffs is very important and interesting because it largely determines the value of many risky assets. Any investor should be compensated for this uncertainty, and valuing assets or investments is a challenging task. Investments are not always only about generating good returns. Suppose you invested in an equity portfolio which generated an annual return of 8%. The same year you observe that your competitor generated an annual return of 15%. This might not sound too satisfying for you, but you have to take into account the risks associated with each investment to make an accurate assessment. If the signs in front of the figures were reversed, then suddenly your portfolio becomes more attractive relative to your competitor.

In modern times, factor models have become a popular tool to assess why different assets exhibit different returns. A factor model is an econometric model which decomposes the influential forces on asset returns into multiple common factors in a linear fashion. The most popular factor models in academic research are the Fama and French [1993] three-factor model, the Carhart [1997] four-factor model, and the Fama and French [2015] five-factor model. In these models, the factors are excess returns on carefully constructed portfolios. In theory, factors should be variables that proxy for effects on investors' utility functions (more on this in section 3.7). There is an infinite number of possibilities when it comes to factor models, and many, many different factors have been hunted for. Many have been successful and many have also critizised the approach reasearchers use to determine factors. In chapter 6, this critique is explained and a discussion about it is elaborated upon. There is no clear-cut consensus to this debate. Factor models are not perfect, however that does not mean that they are useless either. Important insights can still be salvaged from academic reasearch in the field. In practice, these insights are also used by professionals. The biggest investment advisors in the world such as Vanguard and iShares offer plenty of products that stem from factor model reasearch.

Stocks are particularly interesting as an asset class because the stock markets are incredibly popular and enormous, and because many investors have long term investment needs. With long term needs, one has appetite for temporary risks as long as the payoff in the long run is sufficiently high. Just imagine a pension fund that manages investments for customers with target payoffs in, say, 40 years when they retire. The customer's savings are on lockdown, and they can tolerate temporary downturns as long as they get compensated for the risk. Working out how much to compensate is the main challenge. Of course pension funds and other financial institutions face many other challenges such as interest rate hedging, liquidity risk, currency risk, property risk, operational risk, county risk and political risk. In this thesis, however, market risk of stocks will be the main focus.

1.2 Academic Framework

In this section, the foundation and academic scene is set in order to introduce important terms to understand the main idea and meaning of latter sections. The Capital Asset Pricing Model (CAPM) of Sharpe [1964], Lintner [1965], and Mossin [1966] postulates that the expected excess return of a stock is proportional to its market beta, i.e. a measure of systematic risk of the stock in comparison to the market as a whole. The CAPM can be tested because the market return and market beta is something observable or at least something that can be approximated quite precisely. Later in time, Ross [1976] had the idea that the market beta might not be the only systematic risk factor. His Arbitrage Pricing Theory (APT) says that the expected excess return of a stock has a linear relationship with an undetermined number of common risk factors. The fact that the APT does not specify the actual risk factors makes it a more general theory and also makes it impossible to test. The CAPM and the APT both say that all idiosyncratic risk can be diversified and should thus be equal to zero. Alpha (α) is commonly used to denote the excess return of a security or portfolio adjusted for systematic risk.

The CAPM can be viewed as a special and simple case of APT and using the CAPM, the alpha of asset i can be defined as

$$\alpha_i = (R_i - R_f) - \beta_i(R_m - R_f) \tag{1.1}$$

In this equation, R_i is the return of asset i, R_m is the return of the market portfolio, R_f is the risk free rate, and β_i is the ratio of covariance between R_i and R_m to variance of R_m. In a perfect world where markets are perfectly efficient and the model accounts for all systematic risk, it should be impossible to generate any alpha. In the literature, it has become clear that the CAPM is not the best description of influences on stock returns, but if the right factors that capture as much systematic risk as possible are added to the model, alpha should be small in an economic sense.

1.3 Main Research Question

Having introduced the main context and setup of this thesis, the central research question is presented in this section. The central question is inspired by other academic research and practical use of benchmark asset pricing models, and it is as follows:

Are current benchmark factor pricing models able to explain variations in average returns across securities in the sphere of German stocks on German stock exchanges? Which factor pricing models are best able to explain these variations?

The models that are proposed as current benchmark factor pricing models are the Fama and French [1993] three-factor model, the Carhart [1997] four-factor model, and the Fama and French [2015] five-factor model. In this study, a fourth candidate is added which combines Carhart's momentum factor with the five-factor model. Germany is chosen because it has a large enough stock market to find meaningful results, and it has been studied to a less extent than e.g. the US stock market. The research question basically asks whether there is economically and statistically significant alpha in the German stock market. If the proposed models are good descriptions of returns and one or more of them pass all model tests, it would be quite groundbreaking since related literature for other markets maintain that these models are often missing something. It is suspected that there will be significant alpha of some sort, but insights from this study can still be salvaged and have important implications. If significant alpha is found somewhere, it would suggest that the models are inadequate descriptions of the German stock market.

This thesis is related to other academic research. Similar publications include Foye, Mramor, and Pahor [2013], Foye [2017], Blitz [2011], Fama and French [1998, 2012, 2017], Novy-Marx [2013], Aharoni, Grundy, and Zeng [2013], Wang and Yu [2013], Cooper and Priestley [2011], Griffin, Ji, and Martin [2003], Hoel and Mix [2016], Olsen [2016], and Blitz, Hanauer, Vidojevic, and van Vliet [2017].

1.4 Delimitations

In this section, delimiting factors and exclusionary decisions are presented. Delimitations are factors and circumstances that limit and define the scope and perimeters of the study. Delimitations are decisions that are in the author's control and are actively chosen to shape the frame of reference.

There are many different asset classes and types, and new ones appear now and again. Therefore, it is necessary to set a frame of reference. The main focus of this thesis is equity securities, i.e. stocks, mutual funds, or any type of equity portfolio. Academics and theory typically separate bond markets from equity markets for good reasons. Bond markets relate to debt securities and the payoff and risk profiles of these differ quite substantially from equity securities. There are other models and theory that go beyond the scope of this thesis made in attempt to explain bond markets.

In an ideal analysis where all information is available it would be natural to account for transaction costs occurring from securities being traded. Unfortunately the historical data on transaction costs required is not easily accessible. Hence, it would be difficult to estimate transaction costs with an accuracy that is meaningful. Besides, transaction costs are usually very small in comparison to the transaction as a whole. In the most closely related literature to this thesis, many of the authors also limit themselves from transaction costs. While there are some common methods of taking into account some sort of estimated transaction costs, it is decided that transaction costs will not be taken into account in this thesis. The implications of this will be discussed in chapter 6.

Many actors in financial markets are restricted either by their superiors, customers, or moral values from attaining large exposures to short positions in financial securities. Some actors are also restricted from exposures that are too focused on single assets for diversification reasons. Since a decent part of the theory and analysis in this thesis involves long-short portfolios there is a risk that these normal constraints are sometimes exceeded. Another practical problem is that sometimes there are no actors in the other end who is willing to let you take a large short position especially if the securities are small or illiquid. Since the German stock market is relatively large on a world basis, the vast majority of assets in this analysis are highly liquid and most short positions are assumed to be attainable for a sophisticated investor who understands to minimize constraints at least on a small scale. To account for constraints about short exposures would add too much complexity to the analysis to the extent that it would start moving away from the main purpose of the thesis. For these reasons an exclusionary decision is made to assume that there are no constraints in terms of short exposures or too focused exposures to any particular assets.

Multivariate linear regression analysis is the main statistical methodology of choice in this thesis. It is a powerful and widely used method that examines the relationship between multiple independent variables and a dependent variable, however it also has its limitations. Any limitation comes down to the data sample. The data has to live up to the following assumptions: The relationship is linear, the dependent variable is a continous random variable, heteroskedasticity cannot be present, there is zero autocorrelation, and the conditional distributions of the dependent variable are all normal distributions. Certainly, all these assumptions are not always met in all regressions conducted in this thesis. This can lead to imprecise results. The method is used anyway because it would be close to impossible to find practical data that always satisfies these assumptions. The methodology will be further explained in chapter 2.

With professor Kenneth French's website and data [French, 2018], the US stock market has pretty much been scoured for empirical evidence on factors and factor models. A study by Harvey et al. [2016] has shown that more than 300 factors have been found significant to add something to the description of the cross-section of stock returns in the US stock market. They suggest that the hurdle rate for significance is too lenient since so many factors have been found. Obviously it is impossible to consider all possible factors, and there are certainly some aspects of asset pricing that this thesis omits, however, the intention of this thesis is to study the most reasonable and popular factors rather than the whole spectrum. Additionally, rather than studying the US stock market, this thesis will study the German stock market which has also been studied before, but to a far less extent than the US market. This thesis can contribute to the existing literature with an up to date study of a large stock market with newly constructed factors. An empirical study of factor models requires that the subject of study has a certain size for the results to be meaningful. The US market is the biggest in the world, and therefore it has been extensively studied. Germany is chosen in this thesis because it also has several very large stock exchanges. Thus, it is large enough to find meaningful results.

1.5 Data Drive

The following URL links to a drive that contains data files used throughout the analysis in this thesis. The link may be referred to throughout the thesis when relevant in case the interested reader would like to see the data.

https://drive.google.com/drive/folders/1h9o1RIKkmwlznQ-fvuuHXTkFZWIu1Zfp?usp=sharing

Chapter 2

Methodology

This chapter explains the general scientific philosophy and approach that is at hand throughout this thesis. In section 2.1, the research ontology and epistemology is explained and discussed. Section 2.2 explains the methodology that is used in this study. Section 2.3 presents the general empirical methods used in the analysis and explains these methods in detail.

2.1 Philisophy of Science

In finance and economics, there are often theoretical and practical debates about the justification of theories, concepts, or phenomena. In other words how we can know whether something is right or wrong. It is different from e.g. certain areas of physics where scientists make exact measurements and everyone agree that the measurement is correct which tells us something non-debatable about the world. In economic theory, there is a common distinction between the positive view and the normative view. The debate is over general theories and whether they actually tell us how the world works or whether we as social actors deviate from how the world should work.

In order to clarify this discussion, it is first of all necessary to make distinctions between the philosophical terms ontology, epistemology, and methodology. Ontology is the study of the nature of existence. It questions whether different entities exist and how we know they exist. Epistemology is the study of the nature of knowledge or how knowledge is developed. It questions knowledge and asks how we know that we know. Methodology is a more practical term that has to do with the methods used to generate knowledge and to understand the world. In short, ontogology has to do with existence, epistemology has to do with philosophy, and methodology has to do with practice.

In social sciences and business research, ontology can be perceived in two essential aspects, namely whether social entities are objective or subjective. Objectivism is the ontological view that social phenomena exist independently from social actors interested in their existence. In the other end of the spectrum, subjectivism is the view that social phenomena is created from the perceptions and actions of social actors that are concerned with the existence of those phenomena [Bryman, 2012].

Positivism in its extreme form asserts that all knowledge can be verified, and only if knowledge is valid, it is scientific knowledge. Positivists believe that natural phenoma is the foundation of knowledge, and that knowledge is only certain knowledge if it can be experienced or in other words if empirical evidence can be found. Positivism has been under heavy critique and it is said to be pointless since it classifies almost all knowledge as non-certain. For example Werner Heisenberg, who was a theoretical physicist and stood against positivism, said that the positivists have a simple answer to philosophy: "The world must be divided into that which we can say clearly and the rest, which we had better pass over in silence. But can anyone conceive of a more pointless philosophy, seeing that what we can say clearly amounts to next to nothing? If we omitted all that is unclear we would probably be left with completely uninteresting and trivial tautologies." [Heisenberg, 1971].

This critique has lead to another philosophical stance called postpositivism or neopositivism. In this stance, postpositivists acknowledge and accept that background knowledge, theory, and personal values of scientists and researchers may have an impact on the observed. Postpositivist retain the belief that there is an objective truth. By recognizing possible biases, postpositivists pursue objectivity. The true picture can only be known imperfectly or probabilistically. In postpositivism both quantitative and qualitative methods are acceptable [Lindlof and Taylor, 2017]. The philosophy of science in this thesis is closely related to the beliefs of postpositivists.

2.2 Research Methodology

It is important to make some methodological distinctions in order to best explain the methodology of this thesis. Knowledge can be justified a priori or a posteriori. A priori knowledge is independent of experience and based on intuition whereas a posteriori knowledge is known by experience. A way to understand the difference is as follows. If one says that all crows are birds, then one's statement is justified a priori because one does not have to see or experience a crow to know that it is a bird. If one says that all crows are black, one's statement is a posteriori justified because one has seen many crows that are black and therefore believe that all crows are black [Russell, 2010]. This analogy can be converted into the empirical analysis in this thesis. An asset pricing model based on empirical data is in its nature a posteriori justified. In addition, model performance is evaluated using empirical data, and if the model fits the data well, it is a posteriori justified as a good model.

Deduction and induction are processes of reasoning. Deductive reasoning uses one or more premises to reach a certain conclusion by logic. An example of deductive reasoning could be: All humans are mortal, Americans are humans, therefore Americans are mortal. This is also called top-down logic because Americans are a subset of the main set, humans. If the premise for a set is true, then the premise is also true for a subset of the set. Induction or inductive reasoning uses specific events to generalize or extrapolate into a general conclusion. This is also called bottom-up logic. Throughout this thesis, statistical methods are used extensively. Statistical methods use inductive reasoning in nature since these methods are about making inference from the sample to the population from which the sample is drawn. This way of reasoning from the particular to the general is associated with epistemic uncertainty. Epistemic uncertainty is due to what we could know in principle, but do not know in practice. One could in principle know true parameters, but in practice parameters are estimated and measurement error is present in these methods. In general, statistics is the business of interpreting observed data in order to get a better understanding of the real world.

2.3 Empirical Methods

This section explains how to apply methods and techniques which have the purpose of testing model performance of different factor pricing models. A factor pricing model assumes that the variation in asset returns appears from the sensitivity of all returns to some number of common risk factors. These risk factors are each associated with a risk premium. Thus, higher exposures to these risk factors should be compensated by higher expected returns. General notation and inspiration for this section is picked up from Cochrane [2005].

2.3.1 Model Overview

General multi-factor models are used to generate estimates and test statistics in the analysis in chapter 5. Model performance tests are conducted using a two-step approach. The first step involves time series regressions and the second step involves cross-sectional regressions. Time series regressions have the purpose to understand variation over time, and cross-sectional regressions have the purpose to understand average returns in relation to betas. A few definitions are required before diving into the equations. R_t^i is the excess return of asset i or portfolio i at time t in excess of the risk-free rate. α_i is the regression intercept of regression i. β_i is a vector of regression coefficients of regression i. f_t is a vector of common risk factors at time t. ϵ_t^i is the regression residual term of regression i at time t. λ_t is the cross-sectional regression coefficient of cross-sectional regression t. The two regression equations are defined as follows.

$$R_t^i = \alpha_i + \beta_i' f_t + \epsilon_t^i \qquad t = 1, 2...T \quad \text{for each } i \tag{2.1}$$

$$R_t^i = \alpha_t^i + \beta_i' \lambda_t \qquad i = 1, 2, ...N \quad \text{for each } t \tag{2.2}$$

The estimates of α and λ are then the averages over time:

$$\hat{\alpha_i} = \frac{1}{T} \sum_{t=1}^{T} \hat{\alpha_{it}} \;, \qquad \hat{\lambda} = \frac{1}{T} \sum_{t=1}^{T} \hat{\lambda}_t \tag{2.3}$$

This approach is known as the Fama MacBeth procedure and was popularized by Fama and MacBeth [1973]. An advantage of the Fama Macbeth procedure is that it is a computationally easy approach to run cross-sectional regressions and estimate test statistics.

2.3.2 Time Series Regressions

The setup in (2.1) is N test portfolios, K factors, and T time points. In matrix notation it looks as follows.

There are N asset-vectors of length T:

$$\text{Across assets:} \quad \begin{bmatrix} R_1^1 & R_1^2 & \dots & R_1^N \\ R_2^1 & R_2^2 & & \\ \vdots & & \ddots & \\ R_T^1 & & & R_T^N \end{bmatrix}$$

There are K factor-vectors of length T:

$$\text{Across factors:} \quad \begin{bmatrix} f_1^1 & f_1^2 & \dots & f_1^K \\ f_2^1 & f_2^2 & & \\ \vdots & & \ddots & \\ f_T^1 & & & f_T^K \end{bmatrix}$$

With time series observations of the factors and excess returns, a multivariate linear regression of R_t^i on f_t can be used to estimate β_i and $\text{Var}(\epsilon_i)$. The regression equation can be written as,

$$R_t^i = \alpha_i + \beta_i^1 f_t^1 + \cdots + \beta_i^K f_t^K + \epsilon_t^i \tag{2.4}$$

The analysis includes a total of 300 time series regressions. Four different factor models are compared. The first model is the Fama and French [1993] three-factor model which consists of the excess market return (EMR), the small minus big portfolio (SMB), and the high minus low portfolio (HML). The second model is the Carhart [1997] four-factor model which consists of EMR, SMB, HML, and a momentum portfolio (WML). The third model is the Fama and French [2015] five-factor model which consists of EMR, SMB, HML, the robust minus weak portfolio (RMW), and the conservative minus aggressive portfolio (CMA). The fourth and last model is a six-factor model which consists of EMR, SMB, HML, RMW, CMA, and WML. The motivation for why these particular factors are used is given in chapter 3. How exactly these factor-mimicking portfolios are defined and created is explained in chapter 4.

2.3.3 Cross-sectional Regressions

The setup in (2.2) is to use the $N \times K$ beta estimates from (2.1) as explanatory variables. The test portfolios are still used as dependent variables, however they are transposed such that there are T regressions with N observations in each regression. There is a regression at each point in time. In matrix notation it looks as follows:

There are K beta-vectors of length N:

$$\text{betas:} \quad \begin{bmatrix} \beta_1^1 & \beta_1^2 & \dots & \beta_1^K \\ \beta_2^1 & \beta_2^2 & & \\ \vdots & & \ddots & \\ \beta_N^1 & & & \beta_N^K \end{bmatrix}$$

There are T asset-vectors of length N:

$$\text{Across assets:} \quad \begin{bmatrix} R_1^1 & R_1^2 & \dots & R_1^T \\ R_2^1 & R_2^2 & & \\ \vdots & & \ddots & \\ R_N^1 & & & R_N^T \end{bmatrix}$$

The betas are now the independent variables. This introduces an errors-in-variables problem because the betas are pre-estimated. When the factors are also traded assets and in particular when the factors are portfolio excess returns, the errors-in-variables problem hardly makes a difference. It does however matter if the factors are something other than portfolio excess returns [Goyal, 2012].

With cross-sectional observations of the estimated betas and the test portfolios, a multivariate linear regression of R_i^t on β_i can be used to estimate λ_t and α_t. The cross-sectional regression in equation form is,

$$R_t^i = \alpha_t^i + \lambda_t^1 \beta_i^1 + \cdots + \lambda_t^K \beta_i^K \tag{2.5}$$

Cross-sectional regressions are different from time series regressions because they are used to determine why average returns vary across assets rather than over time. In a factor model, the assets that have large exposures to factors with high risk premia, should have high expected returns and vice versa. Cross-sectional regressions estimate the factor risk premia, which investors can use for example to determine the cost of equity of a stock or a portfolio of stocks. The classical example of this is the security market line derived from the CAPM. Using a cross-sectional regression, the security market line is the fitted line that goes through the return-beta scatter plot of different assets or stocks. Stocks with high beta have high expected returns and stocks with low beta have low expected returns. The one-factor approach of the CAPM is translated into a multi-factor approach.

A particular approach to determine factor risk premia is the Fama and MacBeth [1973] regressions. They came up with a practical method covering how to estimate parameters for asset pricing models based on two regression stages. The first stage is to run time series regressions to estimate the betas associated with each pricing factor. In particular, each test portfolio is regressed against the risk factors, and the coefficient beta is the exposure to that particular risk factor. The second stage then involves cross-sectional regressions at each fixed time point. In particular, each test portfolio is regressed against the estimated betas from the first stage regression to determine the risk premium for that particular factor. This results in a time series of risk premia for each risk factor and an error term time series. The factor risk premia are then the averages over time, and the pricing errors are the averages over time.

The main advantages of the Fama-MacBeth approach are that the distribution of the estimated risk premia is independent of the number of stocks, which may vary over time and certainly does. The approach is also easily applicable to unbalanced panel data. Only the returns of stocks that exist in one particular period are used and they could differ from returns of stocks that exist in the next period. Another advantage of the Fama-MacBeth approach is that it accounts for cross-sectional correlation in the pricing errors. Variances of estimates each period are not computed using this approach. Instead, the variance of the average estimates using time series is computed. The last advantage is that the Fama-MacBeth approach can allow for time-varying betas if desired [Goyal, 2012].

A simpler approach to determine risk premia is to use the expected value of the factor time series. No cross-sectional regressions are needed and this approach is more intuitive. This simpler approach does not account for cross-sectional correlation and does not have the same advantages as the Fama-MacBeth approach. The advantage of the simple expected value approach, however, is that measurement error is likely to be much smaller. The Fama-MacBeth approach involves several regressions with estimated parameters. The simple aprroach only involves an average and its associated standard deviation. Both approaches are used in this thesis. In chapter 6, the different results of these two approaches are discussed.

2.3.4 Model Tests

In asset pricing, there are two general approaches for testing models. One can either use time series regressions or cross-sectional regressions. It is possible to combine the two approaches as well. In the literature, both the time series approach and the cross-sectional approach are used, and there is no real consensus about which is better. It is recommended to use both these approaches to control that results are consistent across both methods [Goyal, 2012].

In a one-factor setup where the factor is also a return, i.e. factor-mimicking portfolios, the model

$$E[R^i] = \beta_i E[f] \tag{2.6}$$

can be tested by running OLS time series regressions using (2.1). One can start with a simple test, namely the only implication the model has for the data. The implication is that all intercepts, α_i, should be equal to zero. The α_i generated by the regression are equal to the pricing errors. t-tests are used to determine whether the intercepts are distinguishable from zero. Another important goal is to be able to say something about the joint distribution of the α estimates. If the standard errors of α and β are i.i.d. across time, homoskedastic, and independent of the factors, the overall model can be evaluated using the test statistic given by the asymptotic joint distribution of the intercepts,

$$T\left[1 + \left(\frac{E_T[f]}{\hat{\sigma}(f)}\right)^2\right]^{-1} \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha} \sim \chi^2_N \tag{2.7}$$

where T is the number of time points, N is the number of assets, $E_T[f]$ denotes the sample mean, $\hat{\sigma}(f)$ is the sample variance, $\hat{\alpha}$ is a vector of estimated intercepts, and $\hat{\Sigma}$ denotes the covariance matrix of the residuals $E[\hat{\epsilon}_t\hat{\epsilon}_t{}']$ where $\hat{\epsilon}_t$ is a vector of the residuals.

If the standard errors are also normally distributed, the Gibbons, Ross, and Shanken [1989] (GRS) test is a multivariate test that Fama and French report in their papers on their three- and five-factor models, see Fama and French [1993] and Fama and French [2015]. The GRS test statistic is given by,

$$\frac{T-N-K}{N}\left(1 + E_T[f]'\hat{\Omega}^{-1}E_T[f]\right)^{-1} \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha} \sim F_{N,T-N-K} \tag{2.8}$$

where

$$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^{T}\Big[f_t - E_T[f]\Big]\Big[f_t - E_T[f]\Big]' \tag{2.9}$$

Equation (2.8) is an improvement over equation (2.7) because it allows for sampling variation in the estimate of the covariance matrix, and the F distribution is exact if the errors are normally distributed.

In addition to testing the time-series regression intercepts, it would also be natural to test the cross-sectional pricing errors. In order to test whether pricing errors are jointly indistinguishable from zero, the covariance matrix of the sample pricing errors is estimated from the cross-sectional regression estimates using,

$$\mathrm{cov}(\hat{\alpha}) = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\alpha}_t - \hat{\alpha})(\hat{\alpha}_t - \hat{\alpha})' \tag{2.10}$$

The relevant test statistic which follows a χ^2 distribution with $N - K$ degrees of freedom can then be computed and the zero pricing error hypothesis can be tested by,

$$\hat{\alpha}'\mathrm{cov}(\hat{\alpha})^{-1}\hat{\alpha} \sim \chi^2_{N-K} \tag{2.11}$$

The covariance matrix of the pricing errors is singular. This means that a normal inverse cannot be computed because the determinant is zero, and computing the inverse involves division by the determinant, i.e. division by zero which is not mathematically defined. The Moore-Penrose inverse is a generalized inverse. It is also called a pseudoinverse, and it can be used to compute the inverse of the covariance matrix containing singular values. Singular value decomposition and pseudoinverses is a complicated topic in linar algebra, and explaining it fully is beyond the scope of this thesis. However, an important point here is the fact that an infinite number of inverse covariance matrices could in principle be computed depending on the singular value tolerance. For the analysis at hand in this thesis, a tolerance of 0.1 will be used. The value of the test statistic is quite sensitive to this tolerance. This will be discussed further in chapter 6.

Furthermore, it would be natural to test whether the risk premium of each pricing factor is significant. This is done using a standard t-test, hence an estimate of the sampling errors is needed. In the Fama-MacBeth approach, the variances of a risk premia are estimated by,

$$\sigma^2(\hat{\lambda}) = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\lambda}_t - \hat{\lambda})^2 \tag{2.12}$$

The ratio of the risk premium estimate to the square root of the variance in (2.12) yields a t-statistic which is used in the standard way to determine whether each premium is significant.

Chapter 3

Theory and Literature Review

In this chapter, relevant literature about asset pricing factors and their role in explaining returns is reviewed in order to set the academic scene, put the analysis into context, and identify where this thesis fits with prior research in the field. In addition, relevant asset pricing theory is accounted for. It is assumed that the reader has general knowledge about basic finance, classical finance theory, modern portfolio theory, and a few well-known factor models. Section 3.1 introduces the background for how theory and literature on the topic was developed. Section 3.2 reviews the main literature on the size premium, section 3.3 reviews the main literature on the value premium, sections 3.4, 3.5, and 3.6 review literature about additional factors and discuss the reason and motivation for adding the profitability, investment, and momentum factors to a factor pricing model, and section 3.7 explains fundamental asset pricing theory including theory on factor pricing models and arbitrage pricing theory which is the foundation for why factor models work.

3.1 Background

Finance and economics is not always an exact science like many areas of physics and natural sciences. The great debate between Nobel Prize winners, Eugene Fama and Robert Shiller, is historically important in finance. On the one side, you have Fama who believes in the Efficient Market Hypothesis (EMH) hypothesizing that stock prices reflect all relevant information. On the other side, you have Schiller who claims that the EMH was ”one of the most remarkable errors in the history of economic thought” [Shiller, 1992]. In 1981, Shiller pointed out that ”stock prices move too much to be justified by subsequent changes in dividends" [Shiller, 1981]. After the 1987 crash and the tech bubble, it became widely accepted that stock prices seem to be more volatile than reflected by corporate fundamentals. Fama was, however, still in disagreement and has proposed, together with Kenneth French, the three-factor model of 1993 and recently the five-factor model of 2015. These two poles are often referred to as the rational view (Fama) and the behavioral view (Shiller).

In the latter part of the 20th century, the CAPM of Sharpe [1964] and Lintner [1965] was questioned due to Shiller and empirical evidence of market irrationalities. Instead of basing the model on a single risk factor, the market beta, Stephen Ross suggested the Arbitrage Pricing Theory (APT) which is a framework that bases an asset's return on an unspecified number of common risk factors [Ross, 1976]. It is this theory that lays the foundation of the factor models that have been proposed in recent years; most commonly known are the Fama and French three- and five-factor models.

Already in 1981, Rolf Banz found some empirical contradictions of the CAPM. The strongest contradiction was that the size of market equity adds something to the explanation of the cross-section of asset returns [Banz, 1981]. Another empirical contradiction of the CAPM was found by Stattman [1980] and Rosenberg et al. [1985]. They found that a firm's book value of equity divided by the market value of equity adds to the explanation of asset returns. The two factors, size and value, were suggested by Fama and French [1993] in addition to the market factor. However, the vanilla factors themselves were not included in the model. Instead, they created factor-mimicking portfolios due to statistical assumptions that are best applicable to returns, and also because the APT would only be valid if the factors are also traded assets. The size effect was mimicked by a long-short portfolio that is long in small stocks and short in big stocks. The value effect was mimicked by a long-short portfolio that is long in high-value stocks and short in low-value stocks. Carhart [1997] proposed a four-factor model adding a factor that mimics return momentum. The factor is a long-short portfolio that is long stocks with positive momentum and short stocks with negative momentum. Momentum can be defined in many different ways, e.g. short term or long term. The standard is to use two to twelve months [French, 2018]. Momentum was first documented by Jegadeesh and Titman [1993] who document that past winners outperform past losers over 3 to 12-month holding periods. They are particularly interested in the strategy of selecting stocks based on their performance the last 6 months and holding them for 6 months. In addition, they find that the profitability of these strategies is not due to systematic risk, i.e., they control for beta.

Recently, Fama and French [2015] added two new factors to their model and proposed a five-factor asset pricing model which is now, arguably, the most recent state-of-the-art factor model. The two new factors are related to company fundamentals. One is related to operating profitability and the other is related to the level of aggression of investments made by the firm. The five-factor model does not come without criticism even though it exhibits improved explanatory power. Some concerns are expressed in Blitz, Hanauer, Vidojevic, and van Vliet [2017] who argue that Fama and French should include the, by now empirically established and widely accepted, momentum factor. In addition, they argue that the classic CAPM relation between market beta and return is flat and should thus be excluded from the model. Furthermore, they are concerned with the robustness and the economic rationale behind the two new factors. Rewriting the dividend discount model, as Fama and French [2015] do, is not enough to justify rational risk-based explanations of the two new factors; behavioral explanations seem just as plausible. They do not view this model as a solution to the asset pricing debate, but merely a significant step forward.

Kenneth French now runs a popular website containing a myriad of datasets suited for empirical tests of factor models [French, 2018] primarily on US data since the US stock market is the largest worldwide. In recent times, Fama and French [2017] have also had quite some focus on Asian and European markets. They test their proposed five-factor model in these markets and arrive at findings similar to those for the US market, namely that the five-factor model has difficulties explaining average returns of small stocks that invest aggressively despite low profitability.

3.2 Size

In practice, the size of the market capitalization of a listed firm is a matter of interest for investors. It is a natural intuition to have that the largest sized firms are successful firms because having a large size means that they are part of the most popular stock indices such as the S&P500. These firms are often covered in the media and closely followed by society. This means that for large firms, there exist a lot of information. For small sized firms that are not followed as closely by society, less information is available to investors in a relative sense. Banz [1981] who was the first to find the empirical evidence of the size effect, mentions a conjecture that investors are less inclined to hold small stocks due to a relative lack of information and thus require an extra premium for these stocks. Klein and Bawa [1977] find that if insufficient information about a group of securities, investors will have less interest in holding these securities since there is a risk associated with estimating the parameters of the true return distribution. As described above, it is likely that very small firms generate relatively less information than very large firms and if investors have different amounts of information available, they will tend limit their diversification to those groups of securities with the most available information. The small stocks become undesirable to investors. This behavioral explanation for the size effect is merely a conjecture, and it is consistent with empirical evidence from the 20th century. Interestingly, Schwert [2003] and Goyal [2012] find that the size effect has shrunk or even disappeared in empirical samples that include periods after the publication of Banz [1981] and other publications that highlighted the size effect. Schwert [2003] argue that at about the same time of the publication of papers about the size effect, practitioners began to implement the strategies implied by academic research through investment vehicles. Thus, publication of academic papers in empirical finance could be argued to play an important role in investment behavior and therefore also in the variation of stock returns.

Banz [1981] has no real theoretical explanation for the size effect. Banz argues that since the existence of the size effect is unclear, it should be interpreted with caution. One has to be careful about using the size effect as a basis for general theory. It could be that size is just a proxy for one or several factors correlated with size. Fama and French [1993] do not provide any explicit economic foundation for the size effect. They acknowledge that the effect is there and reference other literature. Later on in Fama and French [1995] they study how size and book-to-market ratio is related to earnings and returns. In that paper, they find that the size factor in earnings help explain the size factor in returns. Small firms with low earnings continue to have low earnings and the size effect could thus be explained by a risk premium of persistence in low earnings. However, they have issues with noise in the data used. They acknowledge that they have no convincing evidence and that behaviorists might have other suggestions to explain the effect. Another attempt to explain the size effect by Fama and French [1996] is that investors need a premium for small firms due to financial distress. They build on Chan and Chen [1991] who argue that small value firms have lost market value because their performance has been poor. Their production has been inefficient, and small firms are more likely to have high financial leverage and cash flow issues than large firms. Chan and Chen [1991] call small firms "Marginal firms" because they are more sensitive to recessions in the general economy, and when the economy is in recession, investors would appreciate high returns more, and they would therefore require a premium on small stocks that are more sensitive.

3.3 Value

Another subject of interest to investors is the degree of value of a stock. Investors have two basic approaches in stock selection, namely to pick value stocks or growth stocks. A value stock is a stock that trades at a low price per share in relation to what its company fundamentals may indicate in terms of performance. Thus, a value stock is undervalued in the view of the investor and thus attractive to select. It is often measured by the price-to-book ratio or equivalently the book-to-market ratio of the book value of equity to the market value of equity.

Stattman [1980] and Rosenberg, Reid, and Lanstein [1985] found persuasive evidence of the anomaly that stocks with a high book-to-market ratio (value stocks) outperforms stocks with a low book-to-market ratio (growth stocks) on average. Inspired by these findings, Fama and French [1993] included the HML factor in their three-factor model. Fama and French [1993] did not provide a theoretical explanation for why the value effect makes the factor model perform well. Similar to the attempt to explain the size effect, Fama and French [1995, 1996] also study the value premium. In terms of persistence in low earnings, the results are not convincing. Fama and French [1995] only find that the market and size factors in earnings help explain the market and size factors in returns, but that is not the case with the value factor. Liew and Vassalou [2000] study whether the size and value effects can be risk factors that predict economic growth. They show that these factors are highly correlated with future growth in GDP. This means that they are also highly correlated with future investment opportunities and future consumption. A value firm typically has more tangible assets, and in a recession, the value firm will be inefficient due to excess capacity. Growth firms are more flexible because they can more easily defer investment opportunities to better economic periods. Hence, value firms are pro-cyclic than growth firms, and the cyclical risk make investors require a higher expected return. Pro-cyclic means that the value of the firm is positively correlated with indicators of the overall state of the economy. Carlson, Fisher, and Giammarino [2004] suggest an explanation of the book-to-market effect that is driven by operating leverage. They argue that when demand for a firm's product decreases, the market value of equity decreases relative to book value of equity. If fixed operating costs are increasing in the size of capital stock, the decreasing demand results in higher operating leverage and thus higher risk. Zhang [2005] also studies the value premium and finds that high value firms have difficulties cutting their assets in recession periods. They are less flexible in terms of costly excess capacity than growth firms causing them to be riskier in recession times when the price of risk is high because investors would appreciate high expected returns more in these periods.

3.4 Profitability

Fama and French [2015] add the profitability factor to their factor pricing model. The addition of this factor was motivated by Novy-Marx [2013] who study the predictive power of gross profitability on the cross-section of asset returns. Earlier on, Fama and French [2006] find that lagged earnings as a proxy for expected profitability is related to average returns. Novy-Marx finds that the ratio of gross profit to assets has more explanatory power than lagged earnings. Even though Novy-Marx suggest gross profits, Fama and French [2015] use the ratio of operating profit to total assets. Recently, Foye [2017] finds that gross profitability provides a better description of returns than operating profitability in his United Kingdom sample. Fama and French [2015] argue that profitability adds to the description of average returns provided by the book-to-market ratio using the dividend discount model originally published by Gordon [1959]. This model states that the market value of a stock is equal to the discounted value of expected future dividends discounted with the internal rate of return on expected future dividends. When two stocks have similar expected future dividends but different stock prices it must mean that the stock with the higher price has more risky expected future dividends given that the two stocks are rationally priced. It is then shown by Miller and Modigliani [1961] that the total market value of equity implied by the dividend discount model can be extracted and is equal to the discounted expected value of total future equity earnings less the change in total book value of equity. Fama and French [2015] then argue that in their rewritten dividend discount model, higher expected earnings imply a higher expected return and thus relates profitability to the description of average returns. Wang and Yu [2013] dissects the profitability premium and find that the majority of the profitability premium is found in the negative alpha of firms with low profitability. Overpricing is more frequent than underpricing due to limitations of short selling. They introduce a behavioral suggestion that assets are mispriced. In particular, they think the most plausible explanation for the profitability premium is that market participants underreact to overpricing due to inattentive behavior. Sun, Wei, and Xie [2014] finds that profitability is related to investment frictions. In the US, there are lower restrictions on short selling than in most other countries. They analyze the gross profitability effect in more than 40 countries and find that in countries with high investment frictions, the profitability effect is lower. In China, where friction forces are high, the profitability effect is weak compared to countries where investment frictions are not as high. They argue for the rational investment-based explanation rather than a behavioral explanation of mispricing.

3.5 Investment

The investment effect was documented by Aharoni, Grundy, and Zeng [2013]. They find a negative relation between investment and average return. They validate Fama and French's earlier predictions at the firm level - not at a per share level. Similar to the profitability effect, Fama and French [2015] use their rewritten dividend discount model to argue that higher expected growth in book equity implies lower expected return, i.e. a negative relation. Titman, Wei, and Xie [2004] study the hypothesis that investors tend to underreact to empire building implications of increased investment expenditures. They observe a negative abnormal relation between capital investment and return that is strongest for firms that have higher discretion in terms of investments. In particular, firms with low debt ratios and high cash flows. Firms that markedly increase capital investments exhibit lower returns than more conservative firms. Cooper and Priestley [2011] study investment and asset growth factors, and they find that investment and asset growth can predict macroeconomic development. Their results are highly consistent with rational explanations of the negative relation between investment and returns. Flexible growth firms can defer investment opportunities and easily scale up or down depending on the macroeconomic state. Large firms that are not as flexible are more sensitive to economic downturns. This justifies the rational explanation of the investment effect.

3.6 Momentum

In practice, momentum investing is a popular concept and many market participants utilize momentum strategies in some way or another. The basic idea of momentum is that if a certain trend is discovered, then you expect that trend to continue in a similar fashion in the future. The only thing needed to create a momentum strategy is knowledge about past security prices. This makes momentum an easy strategy to implement, and researchers and market players often view momentum as an indication of market inefficiency. Past security returns and prices are also used in technical analysis which some people believe in and some people are akin to. Griffin, Ji, and Martin [2003] mention that the momentum effect is strong in European markets, less strong in emerging markets, and present in many Asian markets. Furthermore, they review possible explanations of momentum including data mining, behavioral explanations, and risk. They argue that data mining is an unlikely explanation since momentum has accumulated and maintained its presence since the original study by Jegadeesh and Titman [1993]. Most behavioral explanations have to do with the possibility that when investors obtain new information, they form and review expectations imperfectly based on this information. Chan, Jegadeesh, and Lakonishok [1996] find that earnings momentum exist together with return momentum. Lee and Swaminathan [2000] apparently find that firms with high turnover are more commonly subject to momentum effects. Hong, Lim, and Stein [2000] document that momentum is more present in small sized firms which are less frequently covered by analysts.

Hvidkjaer [2006] studies momentum portfolios and emphasizes that small traders underreact or react more slowly to past returns than large traders. He links momentum to liquidity issues and suggests that the behavior of small traders partly drive the momentum effect. Avramov, Chordia, Jostova, and Philipov [2007] document that the momentum effect is strong in firms with poor credit ratings. These explanations link momentum to liquity risk and credit risk, which are both economically sound and intuitive risk aspects. Hence, they belong with the rationalists.

Humans are often subject to a cognitive bias where they underpredict the amount of variation in a sample. This is known as the clustering illusion in psychology. We look for patterns where patterns might not be. It is fundamentally engraved in us to systematize things in attempts to find meaning. When we only see a tiny part of the world, we desperately want to make sense of the rest. We connect the dots, fill in the gaps with what we think we know, and we update our mental view of the world. We associate systematic patterns with positive feelings and we have negative feelings associated with chaos. Part of momentum investing may simply be due to our human nature and how our brains work. This bias make investors absorb and interpret new information imperfectly. Fama and French

[2015] do not include a momentum factor in their five-factor model, and they do not provide a pleasing explanation for why momentum is not included. Blitz, Hanauer, Vidojevic, and van Vliet [2017] criticize them and argue that the momentum effect is widely accepted in the finance community by now.

3.7 Asset Pricing Theory

This section describes the fundamental theories of asset pricing and how factor pricing models are usually assessed and tested. The Arbitrage Pricing Theory is also described and exemplified in this section.

3.7.1 The Basic Asset Pricing Formula

The theory in this sub-section is based on, and inspired by, Cochrane [2005] chapter 1. Cochrane is an expert and has been doing research in the field for decades, and his book, Asset Pricing, is an award winning read.

If an investor has a utility function $u(c_t) + \beta E_t[u(c_{t+1})]$, then the value of payoff x_{t+1} from any asset can be determined by the investor's first order conditions, that is the basic consumption-based asset pricing formula:

$$p_t = E_t\left[\beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1}\right] \tag{3.1}$$

where c_t is consumption at time t, and β is a so-called subjective discount factor. All asset pricing and portfolio theory is in priciple born from this equation.

To show that this is true, suppose that an investor can buy as much as he wants of the payoff x_{t+1} at the price p_t. Let e be the level of consumption if none of the asset is bought, and let ξ denote the amount of the asset that is bought. Then the investor is faced with a maximization problem

$$\max_{\xi} \left[u(c_t) + E_t\left[\beta u(c_{t+1})\right]\right]$$

subject to the constraints

$$(1)\;\; c_t = e_t - p_t\xi \qquad\qquad (2)\;\; c_{t+1} = e_{t+1} - x_{t+1}\xi$$

If (1) and (2) are substituted into the objective function and the derivative with respect to ξ is set to zero, the result becomes

$$p_t u'(c_t) = E_t\big[\beta u'(c_{t+1}) x_{t+1}\big] \tag{3.2}$$

which is just another way of writing (3.1). The way to understand (3.2) is to think of $p_t u'(c_t)$ as the marginal loss of utility from buying more of the asset and $E_t[\beta u'(c_{t+1}) x_{t+1}]$ as the marginal expected discounted gain of utility from getting the extra payoff in the future.

Let m_{t+1} denote $\beta \frac{u'(c_{t+1})}{u'(c_t)}$. Then (3.1) can be written as $p_t = E_t\big[m_{t+1} x_{t+1}\big]$. If the asset is a stock, the payoff x_{t+1} is substituted by the gross return R_{t+1}. Omitting the subscripts, the equation simplifies to $1 = E\big[mR\big]$ since return can be thought of as a payoff with price 1. This can be decomposed using the covariance decomposition,

$$1 = E[m]E[R] + \text{cov}(m, R) \tag{3.3}$$

The risk-free rate, R_f, is related to m. In particular, it is equal to the reciprocal of the expected value of m. Using this fact, one can write,

$$E[R] - R_f = -R_f \text{cov}(m, R) \tag{3.4}$$

and this can be further broken up into

$$E[R] - R_f = -\frac{\text{cov}\Big(u'(c_{t+1}), R_{t+1}\Big)}{E\Big[u'(c_{t+1})\Big]} \tag{3.5}$$

(3.5) looks similar to the CAPM relation which is known from classical finance. It basically says that all assets should in expectation be equal to the risk-free rate plus a risk premium which is related to the covariance, or correlation, between consumption and the asset. This can be transformed and written as a return beta representation:

$$E[R_i] = R_f + \left(\frac{\text{cov}(R_i, m)}{\text{var}(m)}\right)\left(-\frac{\text{var}(m)}{E[m]}\right) \tag{3.6}$$

or

$$E[R_i] = R_f + \beta_{i,m}\lambda_m \tag{3.7}$$

In (3.7), $\beta_{i,m}$ is the regression coefficient of R_i regressed on m, and λ_m is the risk premium or price of risk. The i subscript indicates asset specificality. Thus, there is a β for each asset, but the risk premium λ is common for all assets. The risk premium is not necessarily the excess market return. It can be the premium of any risk factor, and there can be indefinitely many risk premia.

3.7.2 Factor Pricing Models

Many empirical studies in asset pricing is conducted and written using a return beta representation. A factor pricing model can in general be written as,

$$E[R^i] = \gamma + \beta_{i,a}\lambda_a + \beta_{i,b}\lambda_b + \beta_{i,c}\lambda_c + ... \quad i = 1, 2, ..., N \tag{3.8}$$

The intercept γ has no subscript, which means that the model is equivalent to saying that all assets should have the same intercept. In the usual model, the factors are excess returns because then the risk premium is equal to the expected value of the factor, i.e. $\lambda_k = E[f^k]$. Usually, the test assets are also excess returns, and if that is the case, the model says that the intercept should be equal to zero. The betas arise from the time series regressions,

$$E[R^i_t] = \alpha_i + \beta_{i,a}f^a_t + \beta_{i,b}f^b_t + \beta_{i,c}f^c_t + ... + \epsilon^i_t, \quad t = 1, 2, ..., T \tag{3.9}$$

The factors f should be something that mimicks marginal utility growth. Many, many factor have been proposed and tested empirically, and the big challenge is to find the best proxy for marginal utility growth. Equation (3.9) is not meant to predict returns in the future. Rather, it is used to figure out what the exposure is to each of the factors. The risk exposures (β_i) are then used in equation (3.8) as explanatory variables to explain average returns across assets. This is called a cross-sectional relation, and one can use the cross-sectional regression,

$$E[R^i] = \gamma + \beta_{i,a}\lambda_a + \beta_{i,b}\lambda_b + \beta_{i,c}\lambda_c + ... + \alpha_i, \quad i = 1, 2, ..., N \tag{3.10}$$

to estimate risk premia and test the factor pricing model. To pass the test, the α_i at the end should be statistically indistinguishable from zero such that there is no pricing error present in the model. The test could for example be based on the sum of squared pricing error.

The risk-free rate plays a role in factor pricing models as well. Suppose all the betas in (3.8) are zero. Then γ represents the expected risk-free rate of the model. Normally, one would use a different approach to estimate the risk-free rate and already subtract it from risky returns resulting in excess returns. If excess returns are used instead, then γ can be omitted from the model,

$$E[R^i] = \alpha_i + \beta_{i,a}\lambda_a + \beta_{i,b}\lambda_b + \beta_{i,c}\lambda_c + ... \quad i = 1, 2, ..., N \tag{3.11}$$

This is a more common way of seeing the model, where the asset returns are excess returns.

3.7.3 Arbitrage Pricing Theory

The bottom line of the Arbitrage Pricing Theory APT of Ross [1976] is "where there is mean, there must be covariance". The APT is a way to understand why the model structure is characterized by factors.[1]

Formally, the APT states that if a series of returns are formed by a linear factor model,

$$R^i = E[R^i] + \sum_{j=1}^{N} \beta_{ij}\Big(f_j - E[f_j]\Big) + \epsilon^i \tag{3.12}$$

where

$$E[\epsilon^i] = E\Big[\epsilon^i\Big(f_j - E[f_j]\Big)\Big] = 0$$

then there exist an expected return-beta representation $m = a + b'f$ linear in the factors that prices the returns. The APT only works if we know the set of factors exactly and if the factors can be traded. In practice, factor models are not perfect descriptions of any asset's return, hence the expected error terms are not exactly zero. If the factors were a perfect description and any factor could be traded, then by the law of one price, $E[R^i] = R_f + \beta_i'\lambda$, i.e. the expected return of asset i is equal to the risk-free rate plus the transposed array of factor exposures times the array of factor risk premia. Approximate APT using the law of one price are attempts to salvage something from the exact factor pricing theory to an approximation where errors are economically small.

Suppose a factor model with N test portfolios and 2 factors. First, estimates of α, β, and ϵ are computed using

$$R^i_{t+1} = \alpha_i + \beta_{i1}f^1_{t+1} + \beta_{i2}f^2_{t+1} + \epsilon^i_{t+1} \tag{3.13}$$

[1]The theory in this section is largely inspired by lecture notes from a course in Advanced Investments tought by professor Cochrane. They are available at https://faculty.chicagobooth.edu/john.cochrane/teaching

Then, the desired conclusion is that $\alpha = 0$, i.e.:

$$E[R^i] = \beta_1\lambda_1 + \beta_2\lambda_2 = \beta_{i1}E[f^1] + \beta_{i2}E[f^2] \tag{3.14}$$

If $\epsilon = 0$, then it has to follow that $\alpha = 0$, otherwise there is arbitrage. To exploit the arbitrage, one would take a long position in R^i and take a short position in $\beta_{i1}f^1 + \beta_{i2}f^2$ and end up with the portfolio return

$$R^i_{t+1} - \left(\beta_{i1}f^1_{t+1} + \beta_{i2}f^2_{t+1}\right) = \alpha_i \tag{3.15}$$

The conditions for arbitrage are fulfilled. One has obtained a position with zero cost which is not random and has a positive return.

In practice, the market is not always in equilibrium. Imagine an ETF tracking the S&P500 index with no tracking error. The return on the ETF is higher than the return on S&P500. It is only an 'arbitrage' until arbitrageurs exploit the arbitrage and move the price of the fund up or the price of S&P500 down until the market is in equilibrium. This leads to an approximate APT where ϵ is small and α is small. The theory basically says that we should see small α when the variance of the residuals is small, or in other words, when the R^2 of the regression is big, *alpha* should be small. This is often the case with well diversified portfolios, but not with individual stocks. Individual stocks typically have a lower R^2. Thus, APT is useful when testing e.g. portfolio managers or portfolios in general.

Chapter 4

Data

In this chapter, choices of databases, datasets, and samples are discussed. In addition, the process of filtering and 'cleaning' data is explained, since it has important implications for analytical results. This chapter also explains how the factors and test portfolios used in this study are constructed.

4.1 Raw Data

It is important to collect high quality input data for an empirical study to ensure high reliability and validity of the results. It is also important to collect enough data to fulfill statistical requirements and rules of thumb. Statistics is not an exact science - almost all statistical models are associated with a degree of uncertainty and measurement error. Insuffiecient amounts of data can lead to higher statistical uncertainty, and therefore it is important to collect sufficient data. How much data that is required for sufficiency depends on the complexity of the models. For the analysis in this paper there are two main dimensions of raw data points. One dimension is the length of the historical time period. Another dimension is the number of stocks. A sufficient time period should capture several different business cycles. The number of stocks should be sufficient if the whole spectrum of different firm characteristics is captured in the dataset.

A very well-known source of data for researchers interested in factor models is the website of French [2018] which contains many useful datasets that are continuously updated. Kenneth French collects raw data from the Center for Research in Security Prices (CRSP) and from Compustat, a part of S&P Global Inc. Unfortunately, Kenneth French's datasets are mainly suited for research in the US stock market. He does provide some datasets on what he views as the European market, but not data that is specific to Germany or the German stock market and not data in the preferred degree of detail.

While there exist relevant datasets for the German stock market, e.g. Artmann, Finter, Kempf, Koch, and Theissen [2012] and Brückner, Lehmann, Schmidt, and Stehle [2015], they are not nearly as well-known as the Kenneth French datasets. Brückner et al. [2015] examine that datasets offered by seven different providers differ considerably due to underlying assumptions about country-specific institutional settings and quality problems with the underlying databases. Brückner et al. [2015] argue that these datasets should be used with caution. For this reason, and because the existing data providers do not provide all the data needed for this analysis, raw data is collected from the Compustat Global database. S&P Global Inc. is a well-established data provider that is used by many professionals and practitioners including Kenneth French. This database is updated frequently with daily prices, and this allows for the most recent available data.

From the Compustat Global database, two datasets are extracted. One dataset contains daily security prices and number of shares outstanding along with other information. This dataset is used to compute stock returns and market capitalizations of stocks. Another dataset contains accounting data along with other information. This dataset is used to construct the factors and test portfolios that are to be analyzed. A more detailed description of the security and accounting datasets is provided in sections 4.4 and 4.5, respectively. The raw datasets can be downloaded from the data drive accompanying this thesis. The link to the drive is in section 1.5.

4.2 Time Period

Stock return patterns are highly dependent on timing. The economy can be in upturns and downturns, and the market can be a bull-market or a bear-market. The goal of the models analyzed in this thesis is for them to be able to explain return variations in all time periods and all states of the economy. Therefore, it is important to collect and include data from several different business cycles. According to the National Bureau of Economic Research (NBER), there has been at least 11 business cycles since World War Two and the average business cycle has lasted 5-6 years.[1] The time period for the collected raw data in this thesis

[1] `http://www.nber.org/cycles.html`

is 1988-01-01 to 2018-01-01, i.e. 30 years of data. This means that the data includes at least 5 business cycles which is desirable given the availability of data and capacity of computing power. Two major financial crises are included in the raw datasets, namely the 2001 technology bubble and the 2008 financial crisis. During these times, volatility has been greater than its long term average, and this has implications for the analysis and results in this thesis. Financial crises are, however, not a recent development in economies. For instance, all the way back in the 1630's, the Netherlands had the Tulip Mania. Thus, bubbles and crises are natural occurances and it would arguably be wrong to exclude natural occurances from the analysis. Volatility is not constant over time.

4.3 Currency Conversion

In Germany, the current currency is the Euro and most of the data in the raw datasets is denoted in Euro. The Euro was introduced on January 1st 1999. Before the Euro the German currency was the Deutsche Mark (DEM). This means that there is a transition in the data from DEM to Euro from 1998 to 1999. This is a problem when growth over time is computed, e.g. stock returns, because Euro and DEM did not have a one-to-one value relationship at the time of introduction. In fact, one DEM was worth about 0.6 USD, whereas the Euro and USD were approximately the same value at the time. To remedy this currency overlap, all data is converted to USD which has existed for the entire time period. Converting everything to USD is preferable for another reason. USD is the currency of the world. Most national banks and financial players have USD reserves and commodities are commonly quoted in USD, e.g. gold, coffee, or barrels of oil. Thus, converting to USD makes the data comparable to most phenomena in the rest of the world.

The raw data is denoted in five different currencies: USD, DEM, EUR, GBP, and AUD. All time series of conversion rates are extracted from the Quandl database using the Federal Reserve Economic Data "FRED" source. Quandl is a standard add-in to R which is commonly used for statistical programming. Currency time series are available at the companion drive for this thesis. The link to the drive is in section 1.5.

4.4 Security Data

This section explains how the raw security dataset is analysed and treated to be used as input for the models that are objects of analysis. The dataset retrieved from Compustat contains the following variables:

- Daily close price
- Adjustment factor (used to adjust for stock splits and dividends)
- Shares outstanding
- Currency code
- Issue type code (indicates the type of stock - e.g. common or preferred stock)
- Issue id (useful if companies have more than one issue of stock to uniquely identify each company issue)

In addition, the dataset has standard identifying information including company name, date, global company key (gvkey), international stock exchange code, standard industry classification code (sic), and country code.

4.4.1 German Stock Exchanges

In Germany there are 8 main stock exchanges: Hamburg, Xetra, Hannover, Berlin, Düsseldorf, Stuttgart, Frankfurt and Munich. Only stocks issued on these exchanges are kept in the dataset. The reason for excluding stocks traded on foreign exchanges is that foreign rules and regulations which are not comparable to German rules and regulations may apply to those companies. The purpose is to assess effects that are specific to Germany. The international stock exchange codes used in the sample are: 115 Berlin, 149 Düsseldorf, 154 Frankfurt, 163 Hamburg, 165 Hannover, 171 Xetra, 212 Munich, and 257 Stuttgart.

4.4.2 Financial Intermediaries

Researchers and practitioners often distiguish between firms in the financial sector and firms in all other sectors. Financial firms such as banks, insurance companies, and retirement companies have the role as financial intermediares in society. These firms are fundamentally different from operational firms in terms of capital structure. They are also subject to seperate regulations. The Basel Committee sets capital requirements for banks and Solvency II for insurance companies. Banks and insurance companies can have balance totals that are completely out of proportion from those of operational firms. For these reasons, all financial companies are excluded from the sample. In particular, companies that have sic codes from 6000 to 6999, both numbers included, are tossed out of the sample.

4.4.3 Penny Stocks

A penny stock is a stock that trades at a low price per share and has a low market cap. The United States Securities and Exchange Commission (SEC) has a formal definition of penny stocks, that is stocks traded under $5 per share, is not listed on a national exchange, and the definition can include securities of certain private companies that do not have an active trading market.[2] The United Kingdom defines penny stocks as stocks traded below £1 per share, and in the Euro area, a penny stock is typically defined as a stock trading below EUR 1 per share. Some stock exchanges have requirements that listed stocks have to be of a certain size. If a stock's price per share is very low, movements in the stock price can lead to disproportional returns which would be a misleading way to reflect company growth. Small movements in small prices would overestimate the actual development of a firm. In Germany, rules about delistings due to low prices were not introduced until recently where the different exchanges introduced stricter rules to limit fraudulent behavior. Some of the worlds most famous con artists have used penny stocks in their fraud schemes. A good example is Bernard L. Madoff who was a chairman of the board at NASDAQ in the 1990s. He committed fraud for billions of dollars and was given a maximum prison sentence. For these reasons, penny stocks are removed from the sample. The criteria for removing penny stocks in the return sample is a share price below $1 and a market cap below $0.6 million.

4.4.4 Return Frequency

In empirical studies of factor models, it is most common to use monthly return observations. Fama and French [1993] and Fama and French [2015] use monthly observations in their three- and five-factor studies on the US market. In this thesis, returns in the datasets are computed on a monthly basis. For each security each month, the security price observation on the last day of the month is extracted. Because of weekends, holidays, and potential missing observations, the filtering

[2] `https://www.sec.gov/fast-answers/answerspennyhtm.html`

limit is set to the 24th day of the month or higher. In particular, if an observation exist on day 24 or higher in a given month, that observation is kept. Everything below day 24 is discarded.

4.4.5 Adjusting for Dividends and Stock Spilts

The adjustment factor in the raw dataset assumes reinvestment of dividends and takes stock splits into account. It takes the value 1 if no dividends or stock splits have occured. If e.g. there is a 2-to-1 stock split, the adjustment factor takes the value 2 for all observations prior to the stock split since the price will drop by half its original value. This means that simple returns can be computed as

$$r_t^i = \frac{\frac{P_t^i}{A_t^i}}{\frac{P_{t-1}^i}{A_{t-1}^i}} - 1 \tag{4.1}$$

Where P^i is the price of security i and A^i is the adjustment factor of security i.

4.4.6 Adjusting for Issues on Multiple Exchanges

When observing and screening the sample, it has come to attention that companies issue stocks on multiple exchanges. This leads to duplicate values of returns and market caps for each firm. Therefore, returns for each firm are computed as the weighted average returns over all exchanges and stock issue types weighted with the number of shares at the given exchange or stock issue type. Market caps for each firm are the totals of each of the market caps on each of the exchanges and stock issue types.

4.4.7 Outliers and Return Distribution

After all the initial filtering and cleaning steps have been completed, the return distribution is assessed. There is a number of well-known empirical facts about the typical distribution of stock returns. Many statistical models are based on the assumption that returns are normally or log-normally distributed. However, the typical distribution has a higher kurtosis than the normal distribution, i.e. there are more very high or very low observations than there would be if returns were normally distributed. The typical distribution is approximately symmetric which is equivalent to a skewness of zero. It also has a high peak, which means that there are more returns close to the center than there are in the normal distribution. The sample distribution follow these empirical facts, however the minimum and maximum observations are so extreme that there is a risk of outlier bias. An outlier deviates markedly from other observations in the sample. It is important to identify and exclude outliers for two reasons. One, an outlier might indicate bad data, i.e. errors when coding the data and two, an outlier might be due to a random extraordinary event which is unlikely to occur again and would introduce robustness issues. Since the distribution of returns is not normal, a simple approach of removing the 0.5% highest and lowest observations, instead of e.g. a z-score, is sufficient for this sample. Before deletion of outliers, there are 166,180 observations, and after deletion of outliers, there are 164,518 observations. Summary statistics before and after deletion of outliers are compared in table 4.1. As is reported in the table, the maximum observation is completely out of proportion before outliers are removed. As a result, the measures for skewness and kurtosis are incredibly high. After outliers are removed, the returns have more realistic numbers for maximum, skewness, and kurtosis which are much closer to the empirical regularities that were mentioned.

TABLE 4.1: Comparison of sample return distribution characteristics before and after deletion of outliers. *Mean* is the arithmetic average of returns, *Std* is the standard deviation of returns, *Min* and *Max* are the smallest and largest observations, *Skew* is the skewness of returns, and *Kurt* is Pearson's measure of kurtosis. For a standard normal distribution, Pearson's kurtosis equals 3.

Sample summary statistics comparison						
	Mean	*Std*	*Min*	*Max*	*Skew*	*Kurt*
Before	1.09%	51.03%	-98.90%	13,688%	172.19	39,681
After	0.31%	13.94%	-49.11%	86.16%	0.89	7.80

The processing of data described in this section is implemented in RStudio. In appendix A, the programming code is documented. Intermediary data files can be found at the data drive. The link is in section 1.5.

4.5 Accounting Data

This section describes how the raw accounting dataset is analysed and prepared to be used as model input. The dataset from Compustat contains the following variables:

- Total assets

- Total liabilities
- Stockholders equity
- Total revenue
- Cost of goods sold
- Total operating expenses
- Selling, general and administrative expenses
- Total interest expenses
- Deferred taxes and investment tax credit
- Total deferred taxes
- Fiscal year-end month
- Currency code

In addition, the dataset includes standard identifying information similar to the security dataset.

4.5.1 Data quality

The raw accounting dataset contains 15,116 observations from 1,114 different firms. An initial screening step is to remove firms with zero assets because that is a sign of insufficient reporting. Only 0.2% of the observations are deleted in this step. In cases where stockholder equity is missing, stockholder equity is computed as total assets less total liabilities. The book value of equity is computed as stockholder equity plus deferred taxes and investment tax credit. The second screening step is to remove observations with negative or zero book value of equity since they introduce distortion in the profitability measure. In particular, if a firm has a negative operating profit and a negative book value of equity, the profitability measure will be positive because the two have a ratio relationship.

4.5.2 Sorting Variables

As an intermediary step in the construction of test portfolios and factor mimicking portfolios, relevant sorting variables are constructed. A sorting variable can be viewed as a firm characteristic. It is used to break the data into certain quantiles to group firms by the relevant firm characteristic. The sorting variables needed for the models in this study are *size*, *book-to-market ratio*, *investment* (asset growth), *profitability*, and *momentum*. The sorting variables are yearly measures because in the construction of test portfolios and factor-mimicking portfolios, rebalancing occurs on a yearly basis. June is chosen as the month of rebalance because firms need time to submit their accounting figures. Most firms have accounting year-ends in December, which gives most firms six months to submit their figures. The June portfolio of firms with certain characteristics is held for a year and then rebalanced with new firms according to changes in firm characteristics. The characteristics are defined as follows.

Size is the market value of equity in the end of June for each firm each year.

Book-to-market ratio is defined as the ratio of book equity at the end of the firm's accounting year to the market equity at the end of December,

$$\left[\frac{B}{M}\right]_t = \frac{(\text{total equity})_{t-1}}{(\text{market cap})_{Dec(t-1)}} \tag{4.2}$$

Naturally, not all firms have account years that end in December. This leads to differences in the time points between the numerator and the denominator of the book-to-market ratio. The ratio makes perfect sense when the denominator and numerator is at the same point in time, however when the time points differ, the ratio will be slightly imprecise. Since the book value of equity is not known at all times, this is the closest one can get to a correct measure. The implications of these time differences will be discussed in chapter 6.

Operating profitability is defined as total revenue short of operating expenses and interest expenses all divided by the book value of equity; everything is per the end of the firm's accounting year,

$$\text{Profitability}_t = \frac{(\text{revenue} - \text{operating costs} - \text{interest expense})_{t-1}}{\text{total equity}_{t-1}} \tag{4.3}$$

With this measure there are no issues with time gaps since everything in the ratio comes from the same annual accounts.

Investment is defined as the relative change in assets from one accounting year to the next, i.e. a standard growth rate calculation similar to calculation of returns.

$$\text{Investment}_t = \frac{(\text{total assets}_{t-1} - \text{total assets}_{t-2})}{\text{total assets}_{t-2}} \tag{4.4}$$

Momentum is defined as the value, in June, of the total return of the three previous months,

$$\text{Momentum}_{\text{June}} = (1 + r_{\text{May}})(1 + r_{\text{April}})(1 + r_{\text{March}}) - 1 \tag{4.5}$$

4.6 Test Portfolios

Classically, the CAPM has been tested using individual stocks as dependent variables, i.e., one can think of the left hand side of the equation as a stock's expected return or the firm's cost of equity capital. Instead of using individual stocks as dependent variables in model tests, portfolios of stocks are constructed and used to test the models. The motivation to use test portfolios instead of stocks is that portfolios reduce idiosyncratic volatility which makes estimates of factor loadings more precise and therefore also estimates of risk premia more precise [Blume, 1970]. Portfolios are better measured because they have lower residual variance. Over time, individual stock betas are more likely to change due changes in inherent characteristics of an individual stock. Portfolios are likely to have more stable betas. Moreover, actual investors would most likely attach themselves to portfolios instead of individual stocks so portfolios are motivated by the desire to closely mimic actual investor behavior. Most modern empirical tests of asset pricing models use test portfolios because Kenneth French's data website contains test portfolios constructed using numerous different sorting variables. Fama and French [1993] and Fama and French [2015] follow Blume's motivation in their three- and five-factor models. They construct test portfolios in many different ways; univariate sorts, bivariate sorts, and even three-way sorts. Constructing and testing portfolios by all the different sorts is beyond the scope of this thesis. The test portfolios constructed for the model tests in this study are 5×5 bivariate sorts on size and book-to-market, 5×5 bivariate sorts on size and profitability, and 5×5 bivariate sorts on size and investment. This is similar to what Fama and French [2015] do. A 5×5 sort means that there are five quintiles for each of the two variables. Quintile means to divide the sample into fifths. For example, the 20% smallest and least profitable firms make up the first test portfolio in the 5×5 size and profitability sorts. The 20% largest and most profitable firms make up the 25th test portfolio in the 5×5 size and profitability sorts. Three different sorts with 25 portfolio in each means that 75 test portfolios are constructed in total.

Table 4.2 reports the average number of stocks in each test portfolio. In general, there are fewer stocks on average in the more extreme portfolios which was also expected. When using the size and book-to-market sorts, the portfolio of stocks which are both the biggest in size and highest in book-to-market ratio has the fewest number of stocks within it on average. When sorting by size and operating profitability, there are on average less stocks that have the biggest size and at the same time the weakest profitability. Intuitively, it makes sense that this is the case because most firms benefit from economies of scale. When applying the size and investment sorts, the portfolio of firms that are the biggest and at the same time invest most conservatively has the fewest number of stocks on average. When comparing the sorts in panel A and panel B, it is interesting to observe the diagonals. in general, the diagonal that goes from Big-Low to Small-High has more stocks on average, and in panel B, the diagonal from Small-Weak to Big-Robust has more stocks on average. In panel C, the smallest, most conservative firms are in the majority as well as the biggest, most aggressive firms. The test portfolios are analyzed in secion 5.2

TABLE 4.2: Average number of firms in each test portfolio

Panel A: Size and Book-to-market sorts					
	Low	2	3	4	High
Small	12	10	11	16	30
2	11	12	17	19	20
3	15	17	15	18	14
4	18	19	17	15	10
Big	22	22	19	11	6
Panel B: Size and Profitability sorts					
	Weak	2	3	4	Robust
Small	35	16	11	7	10
2	18	20	15	13	13
3	13	17	18	16	15
4	9	15	17	19	19
Big	4	11	18	24	23
Panel C: Size and Investment sorts					
	Conservative	2	3	4	Aggressive
Small	29	15	12	11	11
2	18	17	14	14	15
3	13	16	16	15	19
4	10	16	16	18	18
Big	8	14	20	21	16

Blume's motivation about test portfolios rather than individual stocks has, in recent time, been debated. Ang, Liu, and Schwarz [2017] criticise and claim that this motivation is wrong because aggregating information into portfolios increases the standard errors of $\hat{\alpha}$ and $\hat{\lambda}$. Constructing portfolios has two effects on standard errors:

... forming portfolios shrinks the standard errors of factor loadings, but this has no effect on the efficiency of the risk premium estimate. ... The second effect in forming portfolios is that the cross-sectional variance of the portfolios betas changes compared to the cross-sectional variance of the individual stock betas. Forming portfolios destroys some of the information in the cross-sectional dispersion of beta making the portfolios less efficient. [Ang, Liu, and Schwarz, 2017]

Even though the creation of portfolios are criticized, it has been a popular approach since Fama and MacBeth [1973] and Jensen, Black, and Scholes [1972]. Researchers use it and market participants trade based upon it on a regular basis. The discussion in chapter 6 touches upon the background and approach of creating test portfolios and why it is useful.

4.7 Factor-mimicking Portfolios

Six different factor-mimicking portfolios are constructed using sorting variables. Five of the factors are long-short portfolios and the sixth is the classical market factor. The six portfolios are excess market return (EMR), small minus big (SMB), high minus low (HML), conservative minus aggressive (CMA), robust minus weak (RMW), and winners minus losers (WML).

The breakpoints used to seperate small stocks from big stocks is the median, i.e. the top 50% stocks in terms of size is labeled "big" (B) and the botton 50% stocks in terms of size is labeled "small" (S). With the factors HML, RMW, and CMA, the breakpoints are the 70th percentile and the 30th percentile, e.g. for RMW, the 30% most profitable firms are labeled "robust" (R), the 30% least profitable firms are labeled "weak" (W). The firms in the middle are labeled "neutral" (N). Similar abbreviations are used for HML and CMA. This results in nine small stock portfolios and nine big stock portfolios. For the momentum factor, the median is used to seperate winners from losers.

SMB is constructed the same way as Fama and French has constructed it in their five factor model [Fama and French, 2015]. They define the factor as

$$\text{SMB} = \frac{1}{3}\Bigg[\Big(\frac{1}{3}(\text{SH} + \text{SN} + \text{SL}) - \frac{1}{3}(\text{BH} + \text{BN} + \text{BL})\Big) + \\ \Big(\frac{1}{3}(\text{SR} + \text{SN} + \text{SW}) - \frac{1}{3}(\text{BR} + \text{BN} + \text{BW})\Big) + \\ \Big(\frac{1}{3}(\text{SC} + \text{SN} + \text{SA}) - \frac{1}{3}(\text{BA} + \text{BN} + \text{BC})\Big)\Bigg]. \tag{4.6}$$

SMB is the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios. When constructed in this way, the size effect is largely free of influence from effects of the other three factors.

HML is the average return on the two high book-to-market ratio portfolios minus the average return on the two low book-to-market ratio portfolios. Fama and French define the factor as

$$\text{HML} = \frac{1}{2}\Big(\text{SH} + \text{BH}\Big) - \frac{1}{2}\Big(\text{SL} + \text{BL}\Big). \tag{4.7}$$

RMW and CMA is defined in a similar way largely free of the size effect:

$$\text{RMW} = \frac{1}{2}\Big(\text{SR} + \text{BR}\Big) - \frac{1}{2}\Big(\text{SW} + \text{BW}\Big) \tag{4.8}$$

$$\text{CMA} = \frac{1}{2}\Big(\text{SC} + \text{BC}\Big) - \frac{1}{2}\Big(\text{SA} + \text{BA}\Big) \tag{4.9}$$

The market factor (EMR) is computed by averaging the return on all stocks in the trimmed security sample each month, subtracting the risk-free rate of that month, and weighting the excess returns in terms of market capitalization:

$$\text{EMR} = \sum_{i=1}^{N} R_{i,t} w_{i,t} - R_{f,t-1} \tag{4.10}$$

It is $R_{f,t-1}$ because the risk-free rate is forward-looking (known ex ante), and the market return is backward-looking (known ex post). $w_{i,t}$ is the weight of firm i, i.e. the ratio of its market capitalization to the total market capitalization.

The momentum factor (WML) is computed by taking the return each month on the portfolio of all the winners, i.e. stocks that are above the median in terms of the momentum sorting variable and then subtracting the return on the portfolio of all the losers, i.e. stocks that are below the median in terms of the momentum sorting variable. In other words, it is the winners minus losers long-short portfolio. The momentum portfolio is updated yearly in June as is the case with the other factors.

$$\text{WML} = \text{W} - \text{L} \tag{4.11}$$

The processing of data described in sections 4.5, 4.6, and 4.7 is implemented in RStudio. In appendix B, the programming code is documented. Intermediary data files can be found at the companion data drive. The link is in section 1.5.

Chapter 5

Analysis and Findings

In this chapter, the constructed factors and test portfolios are analyzed. In the first part of the analysis, the characteristics of the factors are investigated. This is done by computing relevant summary statistics and simple test statistics in section 5.1. The correlations between factors are investigated as well. In section 5.2, characteristics of the test portfolios are analyzed and assessed in relation to the factors. In section 5.3, the factor pricing models are evaluated and tested using a number of different types of model tests. In section 5.4, second-stage Fama-MacBeth regressions are analyzed and findings are presented.

5.1 Factors

In this section, summary statistics on the constructed factors as well as correlations between factors are reported and analyzed.

5.1.1 Summary Statistics

In table 5.1, time series averages and standard deviations of the factor-mimicking portfolios are shown. Additionally, a two-sided mean significance t-statistic is shown. If the t statistic is below 1.96, the mean is statistically indistinguishable from zero. Both the market factor and the momentum factor have means that are significantly different from zero. The means of SMB, HML, RMW, and CMA are indistinguishable from zero. The market factor has the highest sharpe ratio of 0.19 followed by WML with a sharpe ratio of 0.12. The sharpe ratio is the ratio between the mean excess return and the standard deviation of the return time series. The standard deviation is an indication of risk. Hence, the sharpe ratio is a measure of return per unit of risk. The market excess return has a markedly higher standard deviation than the other factors, however the mean excess return is much higher than the other factors, thus resulting in the highest sharpe ratio.

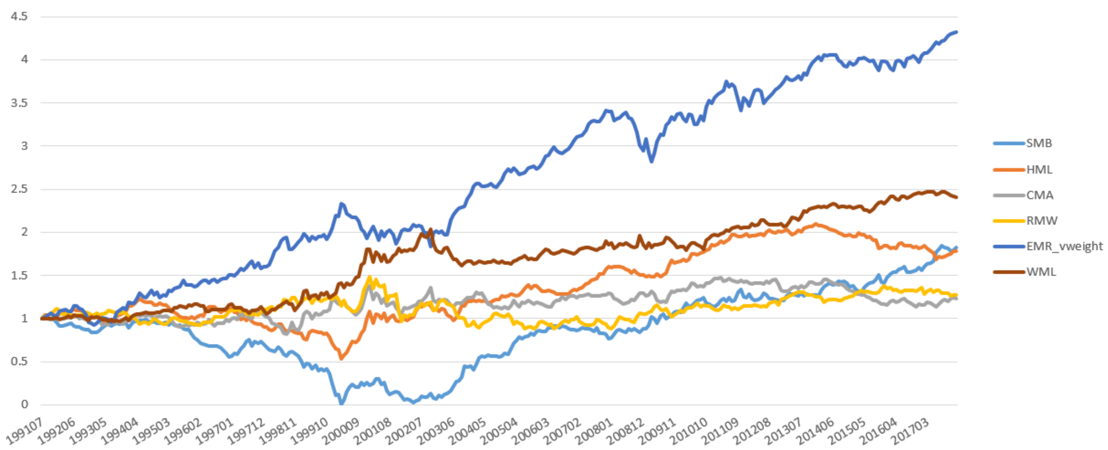
TABLE 5.1: Factor summary statistics in percent. Time series mean, standard deviation, and t-statistic are reported. The statistics are computed using monthly observations from 1991 to 2017. The factors are constructed as described in section 4.7

Summary statistics						
	EMR	SMB	HML	RMW	CMA	WML
Mean	1.04	0.26	0.25	0.09	0.07	0.44
Std	5.47	3.60	3.23	3.67	3.41	3.64
t statistic	3.4	1.28	1.35	0.41	0.38	2.16

In figure 5.1, it is illustrated how the factors develop over time. To make the factors comparable, the figure shows how much one dollar invested in each portfolio at the initial time point is worth at any given future time point. The initial time point is July 1991 and the time series end in December 2017. The graph indicates a high negative correlation between EMR and SMB in the first half of the sample, but then a positive correlation in the second half. This is seen in the gap between the two getting larger roughly during the first 12 years and then staying constant during the remaining time period. The HML factor also appear to be negatively correlated with the market factor in the first quarter or half of the sample and again during the last quarter of the sample. In general, the market factor seems to develop differently from all the other factors. The momentum factor catches up to the market factor in 2003, but after that, the momentum factor appears to be rather flat compared to the market factor which takes off. The implications of the tech-bubble in 2001 and the 2008 financial crisis are seen in the high fluctuations because of increased volatility around these time points.

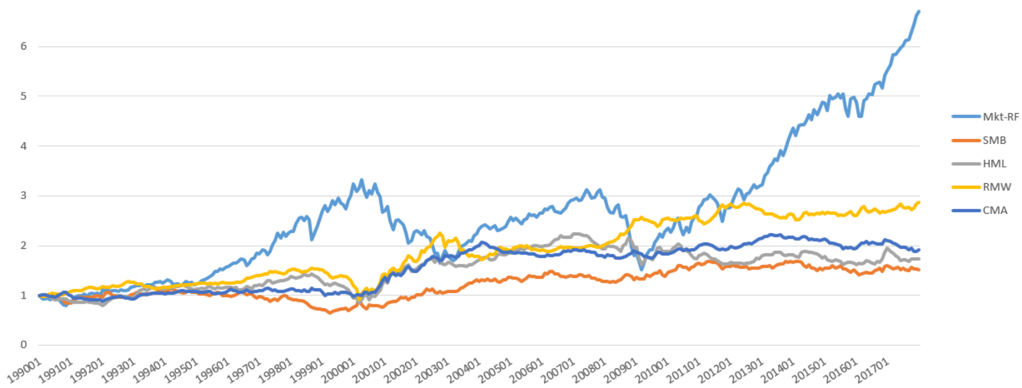
In figure 5.2 the development of the US factors for the same time period is illustrated. The data is obtained from French [2018]. It appears that the US market portfolio has increased a lot in value relative to the German market portfolio in recent years. The SMB factor has been very close to zero in the German market during the years of the dot-com bubble, and in the US market it has not been as low. It has also been the factor with the lowest return up until about 2010 for the German market and throughout the entire time frame for the US market. The RMW factor in the US market appear to have increased much more than it has in the German market. It appears to be close to a 300% return in the US market

FIGURE 5.1: Development of the six factors from 1991 to 2017



while in the German market it has exhibited less than 50% return throughout the entire period. The HML and CMA factors appear have similar developments in the US and Germany.

FIGURE 5.2: Development of the five US-factors from 1991 to 2017. Mkt-Rf corresponds to EMR. WML is not included since it is not used in the US-study.



5.1.2 Correlations

Ideally, a factor pricing model would have explanatory variables that are independent from one another because they should each capture different aspects of the variation in the dependent variable. If independent variables are too similar, the model exhibits multicollinearity issues. As an initial indication of explanatory variable interdependence, the correlations between the factors are shown in table 5.2.

None of the correlations in table 5.2 are above 0.5 and most of the correlations are close to zero. Surprisingly, the correlation between SMB and EMR is -0.49 and

Table 5.2: Correlations between the constructed factors

Correlation matrix						
	EMR	SMB	HML	RMW	CMA	WML
EMR	1	-0.49	-0.04	-0.22	-0.23	-0.16
SMB	-0.49	1	0.03	0.02	0.1	0.02
HML	-0.04	0.03	1	0.12	0.26	0.18
RMW	-0.22	0.02	0.12	1	0.18	0.22
CMA	-0.23	0.1	0.26	0.18	1	0.37
WML	-0.16	0.02	0.18	0.22	0.37	1

the correlation between WML and CMA is 0.37. This is an indication that these explanatory variable pairs have something in common. One possible explanation of the high negative correlation between SMB and EMR is that the market return is value weighted so the high market cap stocks gets a higher weight and SMB involves going short in the stocks with the highest market cap, thus creating returns in opposing directions. The development of the two factors over time is further illustrated in figure 5.1 where this effect can be seen. Also worth noting is the fact that the market factor is negatively correlated with all the other factors. This was also the indication that was given by figure 5.1.

It is also interesting to examine the correlations between the German factors and similar factors for the US market. In table 5.3 the correlations between the factors in the Fama and French [2015] five-factor model and the constructed factors for the German market are reported. The correlation between the German and US market factor of 0.74 is the highest. Apparently there is a negative correlation between the German and US SMB factor, and the other three factors do not appear to be correlated to any significant extent. This means that the constructed factors for the German market are not at all similar to the corresponding factors for the US market created by French [2018]. French also creates factors for other developed markets. To investigate whether factors in non-US markets are correlated with the US factors, table 5.3 also reports the correlations between the ”European” factors and the US factors, the ”Japanese” factors and the US factors, and the ”Asia Pacific” factors and the US factors. These non-US factors are all created by French. When observing these correlations it becomes clear that the factors are completely different depending on which market they are constructed from. This is a bit of a challenge from a theoretical and practical point of view since it would be ideal to have one factor pricing model that works for all markets. It also leads to an excellent discussion about the factors and whether they are only temporary or how they should be interpreted. This is discussed in further detail in chapter 6.

TABLE 5.3: Correlations between the five US factors and the respective constructed factors for Germany and correlations between US factors and non-US factors.

	EMR	SMB	HML	RMW	CMA
Corr(f_{GER} , f_{US})	0.74	-0.16	0.33	0.31	0.17
Corr(f_{EUR} , f_{US})	0.80	0.29	0.58	0.21	0.50
Corr(f_{JAPAN} , f_{US})	0.44	0.07	0.40	-0.05	0.18
Corr(f_{ASIA} , f_{US})	0.71	0.24	0.12	0.20	0.24

5.2 Test Portfolios

In this section, summary statistics and relevant characteristics about test portfolios sorted bivariately by size-book-to-market, size-profitability, and size-investment are analysed.

5.2.1 Size and Book-to-market Sorts

Time series averages, standard deviations, and Sharpe ratios of the 25 portfolios sorted, bivariately, by size and book-to-market ratio are presented in table 5.4.

In panel A, the small-low portfolio, which is the portfolio of small growth firms, has an average return that is markedly above all other portfolios. As shown in table 4.2, the portfolio has an average of 12 constituents which is not much different from the other portfolios. Thus, the high average should not be due to having fewer firms within the portfolio than other portfolios in the sample.

Amongst the smallest firms there is an increasing average return pattern from ”2” to ”High”. This is not the case for any of the other size groups. The pattern of average returns seem arbitrary, and it is difficult to pinpoint any striking relations at first glance. However, the highest value stocks seem to produce average monthly returns in the 1-2 percent range regardless of their size. Additionally, if you dissect panel A into two triangles cutting from small/low to big/high, then the top right triangle returns are generally higher than the bottom left triangle returns. This suggests that small value stocks yield higher returns than big growth stocks.

TABLE 5.4: Size and book-to-market sorted portfolio returns, standard deviations, and Sharpe ratios. The returns are in excess of the risk-free rate. The Sharpe ratio is the ratio of the excess return to the standard deviation.

Panel A: Time series average in %					
	Low	2	3	4	High
Small	2.41	0.87	1.04	1.47	1.70
2	1.20	1.03	1.09	1.40	1.20
3	0.70	1.33	1.24	1.16	1.48
4	0.80	1.21	0.99	1.48	1.98
Big	1.05	0.65	1.17	0.96	1.55
Panel B: Time series standard deviation in %					
	Low	2	3	4	High
Small	9.28	7.98	6.69	6.65	5.80
2	7.05	6.12	5.82	5.65	5.84
3	6.68	5.64	5.36	5.71	6.16
4	5.86	5.60	5.55	5.93	6.99
Big	6.77	5.91	5.82	6.41	7.53
Panel C: Sharpe ratio					
	Low	2	3	4	High
Small	0.26	0.11	0.16	0.22	0.29
2	0.17	0.17	0.19	0.25	0.21
3	0.1	0.24	0.23	0.2	0.24
4	0.14	0.22	0.18	0.25	0.28
Big	0.16	0.11	0.2	0.15	0.21

Value stocks are often considered to have a lower level of risk since value stocks are typically found amongst well-established large firms. In panel B in the size groups "Small" and "2", there is a general decreasing pattern in standard deviation from low to high with one exception in the "2-High" portfolio. However, in the groups "3", "4", and "Big" there is a smiling pattern. In particular, the standard deviation decreases from "Low" to "3" and then increases again from "3" to "High". Thus, only the two smallest size groups are consistent with the notion of lower risk in value stocks.

Small firms are typically more volatile than larger firms. In panel B in the groups "Low", "2" and "3", the volatilities of the two smallest portfolios are lower than the volatility of the three biggest portfolios, so they are consistent with the claim that small stocks are more volatile. However, the pattern in Panel B is a smiling pattern, smiling from small to big in the two highest value groups. Hence, volatility is high in small growth stocks and big value stocks in the sample.

Another thing to note in panel B is the small-low portfolio which has the highest standard deviation. No other portfolios have standard deviations that are close to 9.28%, but for the lowest standard deviation there are other portfolios which are almost as small. This suggests that something strange is happening with the small-low portfolio, and that it might not be robust and apply to other samples.

The Sharpe ratios in panel C tell us something about the efficiency of each portfolio. If rational investors should choose one of these portfolios, they would choose the portfolio with the highest Sharpe ratio because it has the highest return given its level of risk. While the sharpe ratio is not a perfect measure of risk because volatility is not always equal to risk, it still provides initial indications of risk and insights into portfolio efficiency.

The Sharpe ratios in panel C can be compared to the monthly Sharpe ratio of the market portfolio, which is 0.19. The monthly Sharpe ratio for the US market during the same time period is 0.17. Strikingly, the two highest value groups have Sharpe ratios which are all higher than the Sharpe ratio of the German market portfolio with the only exception of the "Big-4" portfolio. In the bottom left corner, the Sharpe ratios are lower than those in the top right corner. This is consistent with the theory that there is a size premium and a value premium, and this speaks in favor of adding SMB and HML to a factor model.

5.2.2 Size and Operating Profitability Sorts

Time series averages, standard deviations, and Sharpe ratios of the 25 portfolios sorted, bivariately, by size and operating profitability are presented in table 5.5.

In panel A, a size pattern is present in the extreme cases. In particular, the smallest portfolios all yield higher returns than the biggest portfolios. This speaks in favor of the size premium and why the SMB factor should be present in factor pricing models.

In the weak-robust dimension in panel A, stocks with the highest operating profitability yield higher returns than stocks with the lowest operating profitability in all size groups except for the smallest group. The second most robust stocks also yield higher returns than the second least robust stocks in all size groups but one. This is an indication that there might be a profitability premium present in the sample. The difference between robust and weak does not appear to be too high, which might suggest that the profitability effect is not significant. In table

Table 5.5: Size and operating profitability sorted returns, standard deviations, and Sharpe ratios. The returns are in excess of the risk-free rate. The Sharpe ratio is the ratio of the excess return to the standard deviation.

Panel A: Time series average in %					
	Weak	2	3	4	Robust
Small	1.77	1.54	1.10	1.79	1.21
2	1.17	1.01	1.07	1.50	1.45
3	1.21	1.11	0.96	1.38	1.34
4	1.14	1.16	1.17	1.12	1.55
Big	0.59	0.92	0.88	1.29	0.89
Panel B: Time series standard deviation in %					
	Weak	2	3	4	Robust
Small	6.62	6.79	6.19	7.14	6.70
2	6.68	6.09	5.38	5.53	5.72
3	8.07	6.04	5.30	5.45	5.32
4	7.79	6.23	5.92	5.09	5.69
Big	9.76	6.56	6.45	6.29	5.86
Panel C: Sharpe ratio					
	Weak	2	3	4	Robust
Small	0.27	0.23	0.18	0.25	0.18
2	0.18	0.17	0.2	0.27	0.25
3	0.15	0.18	0.18	0.25	0.25
4	0.15	0.19	0.2	0.22	0.27
Big	0.06	0.14	0.14	0.21	0.15

5.1 the t-statistic for RMW is not significant which further suggest that the effect is not strong enough. Furthermore, note that the corner with the smallest and most robust firms have higher average returns than the bottom left corner of the biggest and weakest stocks. The lowest average return is in the biggest weakest portfolio.

The highest standard deviation is found in panel B in the biggest and weakest portfolio, and this standard deviation is markedly higher than most other. This is similar to the smallest growth stocks in table 5.4 and might not be robust and apply to other samples. In theory, firms with low operating profitability should be more volatile since they are more likely to experience financial distress. Panel B is consistent with this theory. In the two weakest groups there are no standard deviations below 6%, but in the three groups with the highest operating profitability, the majority has volatility below 6%. The smiling pattern in volatility from small to big is present in all groups except for the weakest group.

The Sharpe ratios in panel C are higher for the smallest portfolios than for the biggest portfolios in all groups. High Sharpe ratios are found in small robust stocks and low Sharpe ratios are found in big weak stocks. Non-extreme size groups in the most robust group have higher Sharpe ratios than the Sharpe ratio of the German market portfolio. The same is true for the second most robust portfolios in all size-groups. Consistent with the pattern in panel A it appears that there is a slight profitability effect that may or may not be economically significant.

5.2.3 Size and Investment Sorts

Time series averages, standard deviations, and Sharpe ratios of the 25 portfolios sorted by size and investment are reported in table 5.6.

Table 5.6: Size and investment sorted returns, standard deviations, and Sharpe ratios. The returns are in excess of the risk-free rate. The Sharpe ratio is the ratio of the excess return to the standard deviation.

Panel A: Time series average in %					
	Conservative	2	3	4	Aggressive
Small	2.06	1.37	1.10	1.09	0.94
2	1.29	0.83	1.26	1.67	0.97
3	1.49	0.89	1.33	1.37	0.95
4	1.16	1.36	1.29	1.13	1.12
Big	1.11	0.97	0.94	0.99	1.18
Panel B: Time series standard deviation in %					
	Conservative	2	3	4	Aggressive
Small	6.62	6.79	6.19	7.14	6.70
2	6.68	6.09	5.38	5.53	5.72
3	8.07	6.04	5.30	5.45	5.32
4	7.79	6.23	5.92	5.09	5.69
Big	9.76	6.56	6.45	6.29	5.86
Panel C: Sharpe ratio					
	Conservative	2	3	4	Aggressive
Small	0.31	0.2	0.18	0.15	0.14
2	0.19	0.14	0.23	0.3	0.17
3	0.18	0.15	0.25	0.25	0.18
4	0.15	0.22	0.22	0.22	0.2
Big	0.11	0.15	0.15	0.16	0.2

In panel A there is a pattern similar to panel A of table 5.5 when it comes to extreme small and big portfolios with an exception in the most aggressive portfolios where the returns are highest for the biggest portfolio and then the average decreases along the big to small dimension.

Panel A suggests that there is a conservative minus aggressive premium present in the sample with the exception of the row with the biggest stocks. The highest average return is found in the smallest most conservative portfolio. When comparing investment group 2 with investment group 4, the investment effect is contradicting in four out of five size groups. In particular, it appears that the second most aggressive stocks yield higher returns than the second most conservative stocks. This is contradicting the CMA effect. Furthermore, the t-statistic of CMA showed non-significance, so it is a bit unclear whether the effect is strong enough or whether the investment effect is even in existence.

In panel B, the pattern of standard deviations is such that where non-extreme size groups cross with the three most aggressive groups, the standard deviations are all low, below 6%. In the conservative groups, the standard deviations are above 6% for all size groups. In theory, firms that invest more aggressively should have higher volatility than conservative firms. Asset growth is a sign of market or product expansion and growth. Panel B is not consistent with this theory, and it is questionable whether growth in total assets is a good measure and a valid candidate in a factor model. The highest standard deviation is found in the biggest most conservative portfolio, and it is markedly higher than all other portfolio standard deviations which suggest that more data is needed to infer anything robust for this portfolio. The high standard deviation might be sample specific.

Peeking at panel C, high Sharpe ratios are found where the three most aggressive portfolios cross with the three middle-sized portfolios. However, the highest Sharpe ratio is found in the smallest most conservative portfolio and, interestingly, this is the winner in the Sharpe ratio category amongst all types of sorts with a Sharpe ratio of 0.31. The standard deviation of this portfolio is 6.62 which is lower than several other portfolios. Thus, it appears as an efficient and attractive portfolio. Worth noting, however, is the fact that this portfolio has 29 constituents on average, which is notably higher than other portfolios. Hence, there is a risk that the data sample in general is too limited and that more data is needed in order to validate results completely.

5.3 Evaluation of Model Performance

This section reports the results of time series regressions. There are four models of interest and for each model, 75 regressions are run because there are 25 portfolios in each of the three bivariate sorting methods. In each regression there are 318 monthly observations, i.e. a period of 26.5 years. The time period is running from the end of July 1991 to the end of December 2017. This results in 300 time series regressions in total. The four models that are evaluated and compared are defined as follows.

Three-factor model:

$$R_t^i = \alpha_i + b_i EMR_t + s_i SMB_t + h_i HML_t, \tag{5.1}$$

Four-factor model:

$$R_t^i = \alpha_i + b_i EMR_t + s_i SMB_t + h_i HML_t + w_i WML_t, \tag{5.2}$$

Five-factor model:

$$R_t^i = \alpha_i + b_i EMR_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t, \tag{5.3}$$

Six-factor model:

$$R_t^i = \alpha_i + b_i EMR_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + w_i WML_t \tag{5.4}$$

Since the right hand sides are not perfect fits for the left hand side, each model also has an error term, ϵ_t^i, which can be added to the equations.

5.3.1 Simple Tests and Explanatory Power

An initial and simple, yet insightful, evaluation of model performance is to determine the rejection frequency of each model. Later, in sub-section 5.3.2, the models will be tested using more sophisticated techniques. The rejection frequency is defined as the number of alphas significantly different from zero over the total number of regressions. First, the rejection frequency of the 25 size and book-to-market, 25 size and operating profitability, and 25 size and investment portfolios are determined seperately for each sorting method, and then finally the total rejection frequency is determined. Furthermore, the adjusted R^2 provides insight about the amount of variation in the dependent variable that is captured by the model. Table 5.7 reports the number of rejected alphas on the 5% and 10% significance levels as well as rejection frequencies and average adjusted R^2. If the intercept is significantly different from zero in a given time series regression, it means that there is a significant amount of variation in the given cross-sectional portfolio that is left unexplained by the model. Thus, any significant intercept would be a sign of weakness for the model.

Table 5.7: Simple test comparison of the four models. The number of rejected alphas on the 5% and 10% significance level as well as rejection frequencies are reported. The last row shows the average adjusted R^2 which is computed by taking the simple average of each of the 25 R^2 for each sort and each model.

Panel A: Size and book-to-market sorts				
Factors	*Three*	*Four*	*Five*	*Six*
# of $\alpha < 0.05$	3	3	3	3
% $\alpha < 0.05$	12.00%	12.00%	12.00%	12.00%
# of $\alpha < 0.1$	5	5	4	5
% $\alpha < 0.1$	20.00%	20.00%	16.00%	20.00%
avg. adj. R^2	63.44%	63.56%	63.88%	63.94%
Panel B: Size and operating profitability sorts				
Factors	*Three*	*Four*	*Five*	*Six*
# of $\alpha < 0.05$	4	3	2	2
% $\alpha < 0.05$	16.00%	12.00%	8.00%	8.00%
# of $\alpha < 0.1$	6	5	8	5
% $\alpha < 0.01$	24.00%	20.00%	32.00%	20.00%
avg. adj. R^2	60.25%	60.50%	62.18%	62.38%
Panel C: Size and investment sorts				
Factors	*Three*	*Four*	*Five*	*Six*
# of $\alpha < 0.05$	3	2	2	2
% $\alpha < 0.05$	12.00%	8.00%	8.00%	8.00%
# of $\alpha < 0.1$	4	6	4	4
% $\alpha < 0.1$	16.00%	24.00%	16.00%	16.00%
avg. adj. R^2	62.65%	63.04%	64.48%	64.69%

Panel A of table 5.7 shows that all models reject 12% of the zero intercept hypotheses in the size and book-to-market sorted test portfolios at the 5% significance level. However, at a 10% significance level, the five-factor model appear to be superior to the other models. The six-factor model has the highest explanatory power measured by the average adjusted R^2, however the difference in explanatory power between the five- and six-factor models is negligible, thus the panel A-winner is likely the five-factor model. In panel B, the five-factor model is inferior to all other models at the 10% significance level. The six-factor model is superior in terms of explanatory power, and it only rejects two test portfolios sorted by size and operating profitability. Hence, the six-factor model is the best performing model in the test portfolios sorted by size and operating profitability. Again in panel C, the six-factor model is the best performing model since it has the highest explanatory power and its rejection frequency at the 5% level and 10% level is 8% and 16%, respectively.

When evaluated at a 5% significance level, the five- and six-factor models appear to have similar overall performance across all sorts. The six-factor model has a slightly higher R^2 than the five-factor model. When evaluated at a 10% significance level, the six-factor model appear to outperform the other models across all sorts. This is, however, only an initial and simple evaluation of model performance, which leaves a bit of insight and a good starting point for further evaluation.

5.3.2 Joint Tests

It is not sufficient to only evaluate models based on individual regressions. It is also relevant to know the joint distribution of the intercepts. The GRS-test follows an F-distribution, and an F-test that all the intercepts are zero simultaneously is used to evaluate model performance.

Recall equation 2.8 which is the GRS equation:

$$\frac{T-N-K}{N}\left(1+E_T[f]'\hat{\Omega}^{-1}E_T[f]\right)^{-1}\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}\sim F_{N,T-N-K}$$

T is the number of time points, N is the number of test portfolios, K is the number of factors. When comparing the models, T and N are fixed across models, however K is different according to which model is tested. Table 5.1 contains the sample means of the factors $E_T[f]$. The matrices $\hat{\Omega}$ and $\hat{\Sigma}$ are the covariance matrices of the factors and the residuals, respectively. $\hat{\alpha}$ is the vector of estimated intercepts.

Table 5.8 reports the GRS test statistics and their associated p-values. It also reports the average of the absolute value of alpha to present an idea of the magnitude of pricing error. The four models are compared across the three different sorts. In most contexts of statistical tests, the goal is to be able to reject the null hypothesis in order to conclude the presence of some relationship between variables. However, in the context of χ^2 and F tests of model fit, the goal is the opposite. In particular, a higher p-value means a better model fit. Equivalently, a lower test statistic means a better model fit. If the GRS statistic was zero, the model would be a perfect fit. What is meant by model fit is how well the model reflects the observed data. Hence, the best performing model, i.e. the model with the least pricing error is the model with the lowest GRS statistic and highest p-value.

TABLE 5.8: GRS test comparison of the four models based on different test portfolio sorts. The GRS statistic is computed using equation 2.8 and the associated p-value is computed using the F-distribution with $N, (T - N - K)$ degrees of freedom. The last row of each panel is the average of the absolute value of alpha.

Panel A: Size and book-to-market sorts				
Factors	*Three*	*Four*	*Five*	*Six*
GRS	1.6869	1.7598	1.6243	1.7324
p-value	0.0235	0.0157	0.0330	0.0183
avg. $\lvert\alpha_i\rvert$	0.0023	0.0023	0.0022	0.0023

Panel B: Size and operating profitability sorts				
Factors	*Three*	*Four*	*Five*	*Six*
GRS	1.4652	1.3424	1.4427	1.4193
p-value	0.0741	0.1312	0.0826	0.0924
avg. $\lvert\alpha_i\rvert$	0.0026	0.0025	0.0022	0.0022

Panel C: Size and investment sorts				
Factors	*Three*	*Four*	*Five*	*Six*
GRS	1.2315	1.2488	1.1817	1.2214
p-value	0.2096	0.1955	0.2544	0.2183
avg. $\lvert\alpha_i\rvert$	0.0018	0.0018	0.0017	0.0018

In panel A of table 5.8 the test portfolios are sorted by size and book-to-market. At a 5% significance level, all models fail the GRS test, but not at a 1% significance level. The best performing model in panel A is the five-factor model since it has the lowest GRS statistic. There are some aspects in the cross-section of firms with different size and book-to-market values that the model fails to capture. To investigate further, recall from table 5.7 that three out of the 25 intercepts were significant. The three portfolios are the "small-low" portfolio, the "3-low" portfolio, and the "big-2" portfolio.

The "small-low" portfolio is the portfolio of the smallest stocks with the lowest book-to-market ratio. Thus, it is an extreme portfolio. It has an average monthly excess return of 2.41% and a standard deviation of 9.28%. The monthly alpha estimated from the five-factor model of this portfolio is 1.14%. These figures are markedly higher than any of the other portfolios.

The "3-low" portfolio is the portfolio of medium sized stocks with the lowest book-to-market ratio. It is characterized by an average monthly excess return of 0.70%, a monthly standard deviation of 6.68%, and a monthly five-factor alpha of -0.51%.

The "big-2" portfolio is the portfolio of big stocks with the second lowest book-to-market ratio. It's average monthly excess return is 0.65% with a standard deviation of 5.91%, and it has a monthly five-factor alpha of -0.40%.

What these three portfolio have in common is that they all have low book-to-market ratios. From that it can be deduced that the model has some difficulties capturing the variation in returns of stocks with low book-to-market ratios. The "small-low" portfolio has a significant positive alpha while the other two portfolios have significant negative alphas. These results may either suggest that there is a lack of good data or that there is some risk-factor associated with these firms that is not included in the model depending on whether you have a positive or a normative view on the results. Either social actors on the market are not acting rationally or there is something missing in the factor pricing model. However, when moving to panel B of table 5.8, the results are more promising. All four models actually pass the GRS test at a 5% significance level. The four-factor model even pass the test at a 10% significance level. Hence, the four-factor model is apparently best at capturing the variation in returns when the test portfolios are sorted by size and operating profitability. In terms of model performance or model fit, the four-factor model is followed by the six-factor model and then the five-factor model. The momentum factor adds something useful to the model since the four-factor model is superior to the three-factor model, which is just the four-factor without momentum. Similarly, the six-factor model is superior to the five-factor model in panel B, and the five-factor model is just the six-factor model without momentum. Recall from table 5.7 that the six-factor model rejected fewer intercepts than the four-factor model. This highlights the point that there is a difference between individual tests and joint tests. Results and findings are more valid and reliable if the analysis is thorough and several aspects of model performance are considered. However, this can also lead to conflicting results as is the case with the four- and six-factor models in this analysis.

In panel C of table 5.8 the results are even more promising. All models pass the GRS test at a 10% significance level, and the five-factor model is the best performer passing the GRS test even at a 25% significance level. The models do not seem to have any problems explaining the size and investment sorted portfolios. Recall from table 5.7 that the five- and six-factor models had the highest R^2 and the fewest rejected intercepts in the size and investment sorted portfolios. Hence, there is conformance between the individual tests and the joint tests in this case.

The results in table 5.8 are quite ambiguous since the five-factor models is the best fit for the panel A sorts and panel C sort while the four factor model is the best fit for the panel B sorts. The next step in the analysis is to evaluate model performance based on the joint test of all 75 regressions at once. This would result in one test statistic and one p-value for each model. Table 5.9 reports the results of these tests including the GRS test statistic, the associated p-value, and the average of the absolute value of alpha for each of the four models. The joint test on all 75 test portfolios shows that the five-factor model outperforms the other models. The five-factor model is the best reflection of the overall data, and it passes the GRS test at a 30% significance level.

TABLE 5.9: GRS test comparison of the four models based on all 75 test portfolios. The GRS statistic is computed using equation 2.8 and the associated p-value is computed using the F-distribution with $N, (T - N - K)$ degrees of freedom. The last row of each panel is the average of the absolute value of alpha.

Based on all 75 test portfolios				
Factors	*Three*	*Four*	*Five*	*Six*
GRS	1.1053	1.1060	1.0907	1.1232
p-value	0.2839	0.2827	0.3089	0.2554
avg. $\lvert\alpha_i\rvert$	0.0023	0.0022	0.0021	0.0021

5.4 Fama-MacBeth Cross-Sectional Regressions

The second stage in the analysis is to determine factor risk premia using Fama-MacBeth second-pass regressions. These regressions are also used to test for overall model pricing error once again with a different point of view. The risk premia are relevant for investors if they e.g. want to determine the cost of equity of a stock or a portfolio using a factor model to price risk. Another way of estimating factor risk premia would be to simply compute the expected value of the factor time series. This was done in table 5.1. Only the market factor and the momentum factor were significant in table 5.1. It would be interesting to compare the Fama-MacBeth risk premia to those found in table 5.1. The improvement of the Fama-MacBeth regressions over the expected values of the factor time series is the fact that the Fama-MacBeth standard errors account for cross-sectional correlation in the error terms. Simple t-statistics do not correct for cross-sectional correlation.

For example, if one firm has an unlikely low alpha, then another (similar) firm also has an unlikely low alpha if they are correlated cross-sectionally. The goal of the cross-sectional regressions is the exposure to each risk premium. The regressions tell us how returns vary in relation to these exposures. The betas from the first-pass regressions are the exposures to the risk premia. Thus, if a coefficient in the cross-sectional regression is significant, it would mean that e.g. a larger exposure to the coefficient's factor leads to a higher expected return, i.e. a risk premium. Recall from equation (2.2) that at each point in time, each test portfolio is regressed against the estimated beta from the regressions in section 5.3. There are 318 time points and 75 test portfolios which means that it is necessary to run 318 second-pass regressions and estimate 636 parameters for each of the four models, i.e. 1272 regressions and 2544 parameters in total. Since this is an extensive amount of data, and since the individual regression parameters are uninteresting to the analysis, only the averages of each parameter are reported.

TABLE 5.10: Cross-sectional regressions to determine risk premia in the five-factor model. The test portfolios are sorted on size and book-to-market, size and operating profitability, and size and investment. Test portfolios are regressed on the factor exposures (estimated betas). The premia and standard errors are estimated using the Fama-MacBeth approach. The t-statistic is the ratio of the premium to the standard error. If the absolute value of the t-statistic exceeds 1.96, the given risk premium is statistically significant.

Panel A: Size and book-to-market sorts					
Factor	*EMR*	*SMB*	*HML*	*RMW*	*CMA*
Premium	0.0100	0.0031	0.0031	-0.0008	0.0123
Std. error	0.0031	0.0025	0.0024	0.0065	0.0057
t-stat.	3.1975	1.2492	1.2764	-0.1225	2.1637

Panel B: Size and operating profitability sorts					
Factor	*EMR*	*SMB*	*HML*	*RMW*	*CMA*
Premium	0.0103	0.0044	-0.0039	0.0051	0.0006
Std. error	0.0031	0.0022	0.0044	0.0027	0.0061
t-stat.	3.2886	1.9920	-0.8770	1.9009	0.1027

Panel C: Size and investment sorts					
Factor	*EMR*	*SMB*	*HML*	*RMW*	*CMA*
Premium	0.0100	0.0021	0.0055	-0.0024	0.0005
Std. error	0.0031	0.0022	0.0047	0.0053	0.0023
t-stat.	3.1681	0.9434	1.1666	-0.4595	0.2057

Table 5.10 reports the Fama-MacBeth risk premia, sampling errors, and t-statistics of the second-pass regressions. The results in the three different panels are mixed.

In panel A, size and book-to-market sorted test portfolio regression estimates are shown. The results are a bit disappointing for the five-factor model because the only significant risk premia are the market risk premium of 1% per month and the investment premium of 1.23% per month. The other three factor risk premia are not even close to being significant. When test portfolios are sorted on size and operating profitability in panel B, the market risk premium is still significant, and then the size premium also becomes significant whereas the investment premium is no longer significant. The profitability premium becomes borderline and the value premium is still not significant. In panel C, the market risk premium is once again a priced factor. None of the other factors show significance. Overall, the results are disappointing because some of the factor risk premia differ according to which characteristic they are sorted by. The only factor risk premium that can confidently be said to be priced is the market risk premium which has a t-statistic above 3, thus a very low p-value. These results are consistent with the previously found results of the factor time-series. In table 5.1, summary statistics of the factor-mimicking portfolios were presented. Recall that the only factors with averages significantly different from zero were the market factor and the momentum factor. The momentum factor is not part of this model, but a resemblance can be seen with the market factor. The results found here are consistent with those found in table 5.1 because the market factor time series average was 1.04% with a t-statistic of 3.4, and here it has the same magnitude and about the same t-statistic and associated p-value.

The other four factors in the five-factor model do not show any clear signs of being priced and this is a concern for the five-factor model. It may be due to the possibility that the data is noisy. The regressions are multivariate regressions with five explanatory variables and the number of observations in each regression is only 25! A general rule in multivariate regression analysis is that at least 10-20 observations per estimated parameter should be maintained. Another explanation for the noisy results could be that the estimated betas may vary over time and the regressions do not account for that. The Fama-MacBeth approach does allow for time-varying betas, e.g. one could use rolling averages as beta estimates instead. Another idea would be to add more observations to each regression. This is done in table 5.11.

In table 5.11 the risk premia, standard deviation, and t-statistic of the second-pass regressions based on all 75 test portfolios are reported. Again the market factor risk premium is 1% per month, has a standard error of 0.31% and a t-statistic of 3.2. None of the other risk premia are particularly close to being significant, however now they are all positive, which we should at least see. These results are still a bit concerning for the five-factor model, however the same points about noisy data as in table 5.10 can be made here. The difference is that the regressions in table 5.11 each contain 75 observations rather than the 25 observations in each of the panels in table 5.10. Table 5.11 is therefore likely to be closer to the actual risk premia since all five of them are positive, and most of them are economically significant (assuming zero frictions to trading) except for the HML risk premium which appear to be very close to zero. If trade frictions were taken into account, all the risk premia would probably be economically small with the exception of the market risk premium. Table 5.11 also reports the test of overall significance of the pricing errors. The p-value of 0.5459 is above the threshold of 0.05, thus the model pricing errors are not too high. The GRS test results in the analysis in section 5.3 indicated that the five-factor model was an adequate model to explain the cross-section of average returns when viewed on a joint basis. The model also passes the joint test of pricing errors in this section. If the model did not pass the joint cross-sectional test, the pricing errors would be too large, and that would be a negative feature for the pricing model. What is negative about the pricing model is the insignificant risk premia. As an ideal, the model would have to pass both joint tests and all risk premia be significant. The test in this section indicates that the pricing error terms are not too large compared to what we should expect, but the risk premia are not large enough compared to their sampling errors.

Table 5.11: Cross-sectional regressions to determine risk premia in the five-factor model based on all 75 test portfolios. Test portfolios are regressed on the factor exposures (estimated betas). The premia and standard errors are estimated using the Fama-MacBeth approach. The t-statistic is the ratio of the premium to the standard error. If the absolute value of the t-statistic exceeds 1.96, the given risk premium is statistically significant. Model fit is the χ^2 test statistic and associated p-value computed using $\hat{\alpha}'\text{cov}(\hat{\alpha})^{-1}\hat{\alpha}$. The number in parenthesis is the test statistic, and to the right of it, the p-value. If the p-value is above 0.05, the model fit is acceptable.

Five-factor model based on all 75 test portfolios					
Factor	*EMR*	*SMB*	*HML*	*RMW*	*CMA*
Premium	0.0100	0.0037	0.0014	0.0044	0.0026
Std. error	0.0031	0.0021	0.0023	0.0027	0.0023
t-stat.	3.2081	1.7338	0.5999	1.6171	1.1134
Joint test $\chi^2_{70}(70.7)$	0.5459				

The analysis in section 5.3 indicated that the five-factor model is the best candidate as a factor pricing model. Therefore it has just been analyzed by itself on a cross-sectional basis where it could be determined that it is not completely without its flaws. Now it is time to compare the five-factor model to the three- four- and six-factor models on a cross-sectional basis. The results of the three- four- and six-factor models using the three sorts with 25 observations in each are not reported because they are just as bad as they are for the five-factor model. Instead, the risk premia and model tests are computed and conducted based on all 75 test portfolios. This adds more observations to each regression making the results more stable.

Table 5.12 reports the risk premia, standard errors, and t-statistics for the six-factor model. The risk premium of the market factor is 1% per month and significant, but none of the risk premia are significant. The joint model fit is a slight improvement over the five-factor model, but that is largely due to the reduction of one degree of freedom. The results look quite similar to the ones in table 5.11. Note here in table 5.12 that the momentum risk premium appear to be very close to zero which is quite disappointing and contradicting with the expected value of the momentum factor in table 5.1 which was significant in terms of its time series t-statistic. Apparently when accounting for cross-sectional correlation using the Fama-MacBeth approach, the momentum risk premium in the six-factor model becomes insignificant.

Table 5.12: Cross-sectional regressions to determine risk premia in the six-factor model based on all 75 test portfolios. Test portfolios are regressed on the factor exposures (estimated betas). The premia and standard errors are estimated using the Fama-MacBeth approach. The t-statistic is the ratio of the premium to the standard error. If the absolute value of the t-statistic exceeds 1.96, the given risk premium is statistically significant. Model fit is the χ^2 test statistic and associated p-value computed using $\hat{\alpha}'\text{cov}(\hat{\alpha})^{-1}\hat{\alpha}$. The number in parenthesis is the test statistic, and to the right of it, the p-value. If the p-value is above 0.05, the model fit is acceptable.

Six-factor model based on all 75 test portfolios						
Factor	*EMR*	*SMB*	*HML*	*RMW*	*CMA*	*WML*
Premium	0.0100	0.0036	0.0015	0.0041	0.0025	0.0015
Std. error	0.0031	0.0021	0.0023	0.0027	0.0023	0.0039
t-stat.	3.2114	1.7046	0.6541	1.5470	1.0881	0.3794
Joint test $\chi^2_{69}(70.6)$	0.5761					

Table 5.13 reports the risk premia, standard errors, and t-statistics for the four-factor model. The momentum factor is included in this model as well, but it does not have a significant Fama-Macbeth risk premium. Neither did the six-factor model indicating that the significance of the t-statistic of the time series mean found in table 5.1 may not be significant when accounting for cross-sectional correlation. The only significant risk premium in the four-factor model is the market risk premium of 1%. The joint test of pricing error is passed with a higher χ^2 statistic than the five- and six-factor models had. Passing this test means that pricing errors are not too high.

TABLE 5.13: Cross-sectional regressions to determine risk premia in the four-factor model based on all 75 test portfolios. Test portfolios are regressed on the factor exposures (estimated betas). The premia and standard errors are estimated using the Fama-MacBeth approach. The t-statistic is the ratio of the premium to the standard error. If the absolute value of the t-statistic exceeds 1.96, the given risk premium is statistically significant. Model fit is the χ^2 test statistic and associated p-value computed using $\hat{\alpha}'\text{cov}(\hat{\alpha})^{-1}\hat{\alpha}$. The number in parenthesis is the test statistic, and to the right of it, the p-value. If the p-value is above 0.05, the model fit is acceptable.

Four-factor model based on all 75 test portfolios				
Factor	*EMR*	*SMB*	*HML*	*WML*
Premium	0.0094	0.0034	0.0031	0.0024
Std. error	0.0031	0.0021	0.0023	0.0038
t-stat.	3.0030	1.6215	1.3359	0.6356
Joint test $\chi^2_{71}(73.5)$	0.6038			

TABLE 5.14: Cross-sectional regressions to determine risk premia in the three-factor model based on all 75 test portfolios. Test portfolios are regressed on the factor exposures (estimated betas). The premia and standard errors are estimated using the Fama-MacBeth approach. The t-statistic is the ratio of the premium to the standard error. If the absolute value of the t-statistic exceeds 1.96, the given risk premium is statistically significant. Model fit is the χ^2 test statistic and associated p-value computed using $\hat{\alpha}'\text{cov}(\hat{\alpha})^{-1}\hat{\alpha}$. The number in parenthesis is the test statistic, and to the right of it, the p-value. If the p-value is above 0.05, the model fit is acceptable.

Three-factor model based on all 75 test portfolios			
Factor	*EMR*	*SMB*	*HML*
Premium	0.0094	0.0035	0.0034
Std. error	0.0031	0.0021	0.0023
t-stat.	3.0108	1.6350	1.4733
Joint test $\chi^2_{72}(74.7)$	0.6094		

Table 5.14 reports the risk premia, standard errors, and t-statistics for the three-factor model. These results show once again that none of the risk premia, other than the market risk premium, are significant, and that pricing errors are acceptable. It can thus be concluded for the four-model comparison that the joint tests of pricing errors are easily passed, and that none of the four models show more than one significant risk premium. These results have to be viewed together with the results of the GRS tests in section 5.2. A simpler approach of estimating risk premia using time series averages of the factors may be sufficient to determine a significant momentum premium, however the results in this section contradicts. Overall it appears that five out of the six analyzed factors all have positive but very small risk premia - so small that they cannot be distinguished from zero. The joint tests of pricing errors are passed for all the models. This joint test statistic is sensitive to the tolerance in the generalized Moore-Penrose inverse of the covariance matrix of pricing errors. When using the RStudio program's standard tolerance, the test statistics become even higher, which does not change the results of the tests. As described in chapter 2, it is recommended by Goyal [2012] to test factor models using both the GRS test and the cross-sectional test and check that results are consistent. Results are consistent. Both tests are passed, so the five- and six-factor models are adequate descriptions of average returns. It can be discussed whether the tests are too lenient, and that is done in chapter 6. Another thing to discuss is statistical parsimony. A parsimonious statistical model is a model that has fewest number of variables or parameters without the cost of explanatory power. In other words, we would like to find the simplest possible model with the greatest possible predictive power. Since the three-factor model also passes the joint tests, it would be a more parsimonious model than the six-factor model because it has fewer explanatory variables, hence it is a simpler model. An even simpler one-factor model consisting of the market risk premium may even be good enough because of its theoretical soundness and statistical parsimony.

Chapter 6

Discussion

In this chapter, the general idea of testing asset pricing models is discussed in terms of how robust, reliable, and valid the models are, and why the interest for these models is so great. In section 6.1, the choice of factors is discussed. Section 6.2 deals with the methodology and how different approaches lead to different results. In section 6.3, the robustness of the results in this thesis is discussed.

6.1 Background and Choice of Factors

Since people in the empirical finance field found out that the CAPM was not without its issues, researchers have eagerly done a lot of asset pricing model tests. The first and most famous researchers to create research portfolios and study how average returns change across different portfolios were Fama and MacBeth [1973] and Jensen, Black, and Scholes [1972] who used individual stock betas to form portfolios based on these betas. Since then, many other average return anomalies have been proposed and studied. Finding these anomalies can be quite easy. Find a variable (e.g. a firm characteristic) that could potentially have something in common with average returns. Assess this variable for all stocks in a given market and sort the stocks into portfolios based on this variable. To determine whether it is an average return anomaly, one can compute betas for the portfolios using regression analysis and test whether the variation in average return is explained by the spread in betas. An anomaly or a 'puzzle' is found if this is not the case.

Some of these firm characteristics or variables may have theoretically sound explanations for why they can explain variations in average returns across assets and some of them may not. The excess market return has a rational and theoretically sound explanation, but variables such as the growth rate of assets or the size of firms may be less economically intuitive. Some of the less intuitive and theoretically sound factors may work very well in one sample, but in another sample, they might not. Recall that the correlations between the factors in the US and other large stock markets were not very high. Ideally, the reason for why factors are priced in the cross-section of assets should be backed by sound theory or intuition. Goyal [2012] argues that there are three main approaches to factor selection. The first approach to select factors uses economic theory and intuition. The Capital Asset Pricing Model is a good example of this approach. The second approach to select factors is statistical. This approach is motivated by the Arbitrage Pricing Theory. Exploratory and confirmatory factor analysis are statistical methods that can be used to analyse the covariance structure of returns in this approach. Another statistical approach is Principal Component Analysis. The third approach to select factors is to find return anomalies based on firm characteristics. This was how the Fama and French [1993] three-factor model began.

In the last 20 to 30 years, empirical researchers have tested an incredibly large number of factors, and basic statistical theory says that if enough factors are tried, some of them will show significance simply due to randomness. A technical thing to note with factor models is that in any historical sample of returns, it is possible to find a portfolio that, when plugged into a factor model, will be able to explain all returns in that sample [Munk, 2016].

The choice of which factors and how many factors to construct in this thesis was made with the above-mentioned things in mind. The size and value factors are obvious candidates because of the popularity of the three-factor model and its well established place in the financial industry. The profitability and investment factors are invented recently and much newer than the size and value factors. This makes them more questionable both in terms of how they are constructed and in terms of whether they have a justified place in a factor pricing model in general. Operating profitability is arguably too far down the income statement, and many people have suggested gross profitability as a better candidate. Investment is originally motivated by growth in equity rather than growth in assets, and whether this makes a significant difference or whether we should use another variable to proxy for investment aggression/conservatism can be discussed and probably needs to be tested more thoroughly. Because of the results in this thesis, it is questionable whether profitability and investment can be used in Germany as priced risk factors.

A more parsimonious model may exclude these two variables because they may be too weak in terms of additional explanatory power.

The momentum factor is an interesting phenomenon to analyze because it can be constructed in many different ways. In this thesis, it was deliberately chosen to use a three-month momentum factor, however it might as well have been based on four, five, six, or even up to twelve months. It is common to describe stock markets as having short term momentum and long term reversal. In particular, over fairly short horizons, e.g. 2-12 months, stock returns exhibit positive autocorrelation, and over fairly long horizons, e.g. 1-5 years, returns are mean reverting. Hence, negative autocorrelation is present. Over very short horizons, e.g. daily, the autocorrelation is typically shown to be positive, but economically non-significant, thus markets are efficient. Algorithmic traders may disagree on this having found certain algorithms that are abnormally profitable, but the focus of the momentum factor is not on these very short horizons. The point of this discussion is that it can be difficult to find a sweet spot where short term momentum is large. It may be two months or it may be close to a year. Three-month momentum was chosen in this thesis based on prior research and intuition. Discussions of autocorrelation, momentum, and mean reversion can be found in Fama and French [1988], Jegadeesh and Titman [1993, 2001], Bondt and Thaler [1985], and Cochrane [2005].

6.2 Methodology

Nearly all studies and analyses have some kind of toolbox that support and lay grounds for that study. In this thesis, an important tool is multivariate regression analysis. There are certain assumptions and rules of thumb associated with these statistical methods. These assumptions are not always easy to fulfill and sometimes the rules of thumb has to be broken. When this happens, the reliability of the output or the results is weakened. Goyal [2012] and Cochrane [2005] survey and review many different approaches and methods that can be used to study asset pricing. In this thesis, only a small number of these methods were chosen due to the scope and resources associated with this thesis. In particular, the Gibbons, Ross, and Shanken [1989] test and the Fama-MacBeth method has been chosen due to a number of advantages of it. One, it is computationally easy and realistic to implement. Two, it accounts for cross-sectional correlation when determining the risk premia. Three, it is useful for panel data. Finally, it implements two different tests of model pricing error that can be compared. The general approach in this thesis is to take advantage of the notion that time series regressions and cross-sectional regressions are complementary, and then use them together.

There are a number of other and possibly better methods that can be used to analyze factor models. Cochrane [2005] reviews several General Method of Moments (GMM) approaches. Furthermore, he reviews several General Least Squares (GLS) methods instead of the Ordinary Least Squares (OLS) methods that have been used in this thesis. The advantage of GMM is robustness of standard errors. Goyal [2012] reviews approaches where factors are non-traded assets and suggests ways to account for the errors-in-variables problem that betas in the second-pass regressions are estimated. He also mentions GMM approaches and how to obtain more robustness in the standard errors. These additional approaches are either too comprehensive, not particularly necessary, or beyond the scope of this thesis.

Another topic of discussion is the way that factors are constructed based on accounting data. Accounting data is available on a yearly basis, but each individual firm have the freedom of choosing when their accounting year ends. These accounting rules impose a challenge in the way that certain accounting figures are compared across firms at a certain point in time. If one firm e.g. has accounting figures that have been measured in June and another firm has accounting figures that have been measured in December, then a comparison of the two firms will be subject to time offsets. Suppose you want to compare the book-to-market ratio of the two firms at a certain point in time. Then you can never compute the book-to-market ratio at exactly the same time for both firms unless you have semi-annual accounting figures. In this thesis the accounting dataset has yearly figures which creates this problem of time offsets in the book-to-market ratio. The best way to remedy this problem would be to collect monthly accounting data, however that kind of data transparency is rarely available.

6.3 Delimitations and Robustness

The analysis in this thesis involves many different long-short portfolios at many different points in time. In practice, it is costly to trade frequently, and it can sometimes be difficult to create certain positions, e.g. high negative exposure. There may be different types of restrictions to certain exposures and positions. These practical facts imply that the analysis in this thesis is not 100% realistic. Some of the portfolios or positions in this thesis may only be possible in theory - not in practice. A deliberate choice not to account for transaction costs and trade restrictions was made before doing the analysis for several reasons. One, these aspects of financial markets would make the analysis highly complex and lead the analysis in other directions and away from its main purpose. Two, the availability and transparancy of transaction costs and trade restrictions is quite poor. Three, the time and resources it would take to obtain meaningful measures of these variables is not worth the minuscule gain in validity and reliability of results. Finally, numerous other researchers that conduct similar studies do not account for these things. Hence, there is no consensus to do so.

The time period that is chosen for the datasets and the analyses in this thesis is important because it is deterministic for the results to a certain degree. The time period from 1991 to 2017 is a relatively recent time period when compared to how long the stock markets in Germany have existed. If older data had been available, this should be included to the datasets, and the results might have looked a bit different. Times are changing all the time, and when new research gets published, people's views on stock markets change. New trading algorithms gets invented and alter the behavior of stock prices to some degree. Firms enter and exit the stock markets from time to time. Markets are ever-dynamic, and these things all factor in to the behavior of asset prices. Therefore, in order to find the best results, one must test the pricing models in many different time periods and collect many different samples. Otherwise, results may only be valid in-sample.

Sometimes there are different ways of estimating or viewing the same thing. In this thesis, risk premia are estimated in two different ways. Using a simple approach gives the result that the momentum premium is significant, but using the Fama-MacBeth approach gives the result that it is not significant. There are also lots of ways to find the inverse covariance matrix for the joint test of pricing errors in the cross-sectional regressions since it is a singular matrix. This means that one cannot be absolutely certain about everything. The point of statistics is not to find an exact mathematical result, but rather to infer something insightful from models that make estimates and strive to make those estimates as precise as possible.

Chapter 7

Conclusion

This thesis evaluates various factor asset pricing models and their ability to explain average stock returns of German stocks on German stock markets in a sample period spanning from 1991 to 2017. Six factors and 75 research portfolios are constructed from scratch, and four different factor models are tested and compared. A Fama-MacBeth approach which exploits the idea that time series regressions and cross-sectional regressions are complementary is used. The four proposed models are evaluated using a two-pass regression approach.

The constructed factors relate to covariation with the market, firm size, firm value, firm profitability, investment aggression, and a momentum factor. All factors have positive monthly average returns, however the market factor and the momentum factor are the only factors of which the monthly average returns are statistically significant. This suggests that the market factor and the momentum factor are the only factors with priced risk premia.

The market factor is negatively correlated with the other five factors, however these correlations are weak. Furthermore, the correlations between the constructed factors are in general weak, so multicollinearity is not a concern. The factors are also weakly correlated with similar factors in other stock markets in other regions. This suggests that it is difficult to use the same factors in a factor model regardless of region, country, or stock exchange.

The constructed test portfolios are based upon a double-sorting technique. 25 portfolios sorted on size and book-to-market ratio, 25 portfolios sorted on size and profitability, and 25 portfolios sorted on size and investment aggression are used as test portfolios in the evaluation of the four different factor models.

Time series regressions show that a five-factor model consisting of a market factor, a size factor, a value factor, a profitability factor, and an investment factor does the best job of explaining average returns. A six-factor model which consists of the five mentioned factors plus a momentum factor performs almost as well as the five-factor model in explaining average returns. Both models pass the test that pricing errors are jointly indistinguishable from zero which means that they are adequate explanations of the variation in average returns. A three-factor model consisting of the market factor, size factor, and value factor as well as a four-factor model that consist of those three factors plus the momentum factor are also analyzed, and based on time-series regressions, these models are inferior to the five- and six-factor models. The five-factor model and six-factor model have one main weakness. They are not able to accurately describe firms with low ratios of book equity to market equity. The models are either missing a risk factor related to these firms or the data is poor for firms with low book-to-market ratios.

The four different factor models are also evaluated in terms of cross-sectional regressions and in that evaluation, they all pass the joint tests of pricing errors consistent with the joint tests in the time series regressions. This means that model adequacy is confirmed. The main issue with the four proposed factor models is that the only significant risk premium is the market risk premium. For the other five proposed factors, it cannot be concluded that the risk premia associated with them are different from zero. This suggests that a one-factor model consisting of the market factor may be a more parsimonious description of average returns.

Current benchmark asset pricing models are able to explain variations in average returns across German stocks in German stock markets with minor weaknesses. The Fama and French [2015] five-factor model is the best out of the proposed models. This model is not a perfect description of variations in the average returns, but it captures many aspects of these variations, and it can be useful for benchmarking purposes and, in general, for investment decision making purposes. A universal factor model that is adequate for all segments and all markets worldwide is difficult to create since factors appear to be area specific. The model needs to be calibrated and tailored for a given market. Many ideas to improve the Capital Asset Pricing Model exist. Creating one perfect overall asset pricing model remains a huge challenge and it is difficult to arrive at an agreement or a consensus on this topic in the field of asset pricing.

Appendix A

R-code: Security Data

```
#Setup
    library(data.table)
    SD <- read.csv("C:/Rstudio/germany_raw_sd_yymmdd.csv")

#Removing financial firms
    SD <- subset(SD, sic > 7000 | sic < 6000)

#Renaming
    colnames(SD)[3:4] <- c("Date","company")

#Extracting the last day of each month
    SM <- SD[c(diff(as.numeric(substr(SD\$Date, 9, 10))) < 0, T), ]

#Creating seperate columns for date variables
    SM$Year <- as.numeric(substr(SM$Date, 1, 4))
    SM$Month <- as.numeric(substr(SM$Date, 6, 7))
    SM$Day <- as.numeric(substr(SM$Date, 9, 10))}

#Changing price and shares outstanding to numeric variable types
    SD <- transform(SD,
        cshoc=as.numeric(gsub(",",".",cshoc)),
        prccd=as.numeric(gsub(",",".",prccd)))

#Only stocks traded on German stock exchanges
    SM <- subset(SM, exchg == 115 | exchg == 149 | exchg == 154 |
        exchg == 163 | exchg == 165 | exchg == 171 | exchg == 212 |
            exchg == 257)

#Only common and preferred stocks
    SM <- subset(SM, tpci == 0 | tpci == 1)
```

```
#Only end of month observations
    SM <- subset(SM, Day > 24)

#Subsetting to relevant variables
    SM <- subset(SM, select=c("Year","Month","iid","gvkey","company",
        "curcdd","prccd","cshoc","tpci","ajexdi"))

#One and only one observation per month
    SM <- within(SM, id <- paste(gvkey, iid, sep = "_"))
    SM <- as.data.frame(SM)
    data1 <- unique(SM[,c("Year","Month")])
    data2 <- unique(SM[,c("company","id")])
    data1 <- merge(data1,data2)
    SM <- merge(SM, data1, all=T)

#omitting  missing values
    SM <- na.omit(SM)

###Currency conversion is not done in R. New setup with new file
    GER_M <- read.csv("C:/Rstudio/Germany_m_usd.csv")

#Calculation of returns
    rfunction <- function(r) {c(NA, diff(r) / r [-length(r)])}
    GER_M <- data.table(GER_M)
    GER_M[, RETURN := rfunction(adjp_USD), by = id]
    GER_M <- GER_M[order(company, Year, Month), ]
    GER_M <- GER_M[, list(tShares = sum(cshoc),
        tPrice = weighted.mean(prccd_USD, cshoc),
        RETURN = weighted.mean(RETURN, cshoc)),
        by = c("Year", "Month", "gvkey", "company", "curcdd", "RF")]
    GER_M <- na.omit(GER_M)

#Calculation of market capitalization
    GER_M$ME_USD <- (GER_M$tShares * GER_M$tPrice)
    GER_M$ME_USD_mil <- (GER_M$ME_USD / 1000000)

#Removing penny stocks
    GER_M$penny <- ifelse(GER_M$tPrice < 1 & GER_M$ME_USD_mil < 0.6,
        1, 0)
    GER_M <- subset(GER_M, penny == 0)

#Sorting by company, then by time
    GER_M <- data.table(GER_M)
    GER_M <- GER_M[order(company, Year, Month),]
```

```
#These are the desired variables
    GER_M <- subset(GER_M, select = c("Year", "Month", "gvkey",
    "company", "RETURN", "RF", "ME_USD_mil", "tPrice", "curcdd"))

#Calculate excess return
    GER_M$excess_ret <- (GER_M$RETURN - GER_M$RF)

###New setup after saving file before trimming begins
    library(moments)
    GER_MA <- read.csv("C:/Rstudio/Ger_before_trimming_v2.csv")

#Descriptives before
    rBef <- GER_MA$RETURN
    before <- data.frame(min(rBef, na.rm = T),
        max(rBef, na.rm = T), mean(rBef, na.rm = T),
        sd(rBef, na.rm = T), skewness(rBef, na.rm = T),
        kurtosis(rBef, na.rm = T))
    colnames(before)[1:6] <- c("Min", "Max", "Mean", "Std",
        "Skewness", "Kurtosis")

#Removing outliers
    high <- quantile(GER_MA$RETURN, 0.995, na.rm = T)
    low <- quantile(GER_MA$RETURN, 0.005, na.rm = T)
    GER_MA <- subset(GER_MA, RETURN >= low & RETURN <= high)

#Descriptives after
    rAft <- GER_MA$RETURN
    after <- data.frame( min(rAft, na.rm = T),
        max(rAft, na.rm = T), mean(rAft, na.rm = T),
        sd(rAft, na.rm = T), skewness(rAft, na.rm = T),
        kurtosis(rAft, na.rm = T))
    colnames(after)[1:6] <- c("Min", "Max", "Mean", "Std",
        "Skewness", "Kurtosis")
```

RStudio-code written by Christoffer Iversen with inspiration from Hoel and Mix [2016].

Appendix B

R-code: Accounting Data, Portfolio Construction, and Factors

```
# Setup
    library(data.table)
    AY <- read.csv("C:/Rstudio/germany_raw_fy_yymmdd.csv")

# Renaming
    colnames(AY)[7] <- "Year"
    colnames(AY)[9] <- "Date"
    colnames(AY)[24] <- "company"

# Only companies from German exchanges
    AY <- subset(AY, exchg = 115, exchg = 149, exchg = 154, exchg = 163,
        exchg = 165, exchg = 171, exchg = 212, exchg = 257)

# Seperate balance sheet items and p&l items
    AY_bal <- subset(AY, select = c("Year", "fyr", "gvkey", "company",
        "curcd", "at", "lt", "seq", "txditc"))
    AY_pnl <- subset(AY, select = c("Year", "fyr", "gvkey", "company",
        "curcd", "revt", "cogs", "xsga", "xint", "xopr", "xopro"))

### Currency conversion is not done in R. New setup
    library(data.table)
    AY <- read.csv("Germany_FY_all_usd.csv")

# Toss out firms with zero assets
    AY <-  subset(AY, at_USD != 0)
```

```
# Defining the book value of equity
    AY[, equity := ifelse(!is.na(seq_USD), seq_USD, (at_USD - lt_USD))]
    AY$BE_USD <- (AY$equity + AY$txditc_USD)

# Toss out firms with zero or negative BE
    AY <- subset(AY, !BE_USD <= 0)

# Defining profitability
    AY[, Profitability := (revt_USD - xopr_USD - xint_USD) / BE_USD]

# Time frame
    AY <- as.data.frame(AY)
    data1 <- unique(AY[, "Year", drop = F])
    data2 <- unique(AY[, c("company", "gvkey")])
    data1 <- merge(data1, data2)
    AY <- merge(AY, data1, all = T)

# Defining a function to calculate investment
    rfunction <- function(r) {c(NA, diff(r) / r [-length(r)])}

# Defining investment
    AY <- data.table(AY)
    AY <- AY[order(company, Year),]
    AY[, Investment := rfunction(at_USD), by = company]

# Missing values are omitted
    AY <- na.omit(AY)

# Reading return data
    data1 <- read.csv("Ger_after_trimming_v2.csv")
    data1 <- subset(data1, Month == 12, select = c("Year",
    "company", "ME_USD_mil"))

# Reading BE, Profitability and Investment
    data2 <- subset(AY, select = c("Year", "company", "BE_USD",
        "Profitability", "Investment"))

# Merging
    G4F <- merge(data1, data2, by = c("Year", "company"))
    G4F <- data.table(G4F)
    G4F <- G4F[order(Year, company),]

# Book-to-market ratio
    G4F$BM <- (G4F$BE_USD / G4F$ME_USD_mil)
```

```
    G4F <- subset(G4F, select = c("Year", "company", "BM",
        "Profitability", "Investment"))
    G4F$Year <- (G4F$Year + 1)

# BE, Profitability, and Investment t and ME June t+1
    GER_SM <- read.csv("Ger_after_trimming_v2.csv")
    data3 <- subset(GER_SM, Month == 6, select = c("Year",
        "company", "ME_USD_mil"))
    colnames(data3)[3] <- "size"
    G4F <- merge(G4F, data3, by = c("Year", "company"), all = T)
    G4F <- na.omit(G4F)

### New Setup
    library("plyr")
    library("data.table")
    test_pf <- read.csv("Sorting_variables.csv")

# Adding quintile indicator columns
    test_pf <- ddply(test_pf,.(Year),transform, size_gr = cut(
        size, breaks = c(quantile(size, seq(0,1,by=0.2))),
            labels = c("1", "2", "3", "4", "5"), include.lowest = T))
    test_pf <- ddply(test_pf,.(Year),transform, BM_gr = cut(
        BM, breaks = c(quantile(BM, seq(0,1,by=0.2))),
            labels = c("1", "2", "3", "4", "5"), include.lowest = T))
    test_pf <- ddply(test_pf,.(Year),transform, Inv_gr = cut(
        Inv, breaks = c(quantile(Inv, seq(0,1,by=0.2))),
            labels = c("1", "2", "3", "4", "5"), include.lowest = T))
    test_pf <- ddply(test_pf,.(Year),transform, OP_gr = cut(
        OP, breaks = c(quantile(OP, seq(0,1,by=0.2))),
            labels = c("1", "2", "3", "4", "5"), include.lowest = T))

# Adding double indicator columns
    test_pf <- within(test_pf, size_BM <- paste0(size_gr, BM_gr))
    test_pf <- within(test_pf, size_OP <- paste0(size_gr, OP_gr))
    test_pf <- within(test_pf, size_Inv <- paste0(size_gr, Inv_gr))

# Each observation 12 times
    LHS <- data.frame(test_pf[rep(seq_len(nrow(test_pf)),each = 12),])

# Adding a month column and merging with return data
    LHS$Month <- rep(c(7:12,1:6), times = nrow(LHS)/12)
    LHS <- data.table(LHS)
    LHS[, Year_ret := ifelse(Month > 6, Year, Year + 1)]
    return_data <- read.csv("Ger_after_trimming_v2.csv")
    colnames(return_data)[1] <- "Year_ret"
```

```
    LHS <- merge(LHS, return_data, by = c("Year_ret", "Month", "company"))

# Compute average excess return of each test portfolio
    LHS <- data.table(LHS)
    Size_BM_ret <- LHS[, list(R_eweight = mean(excess_ret),
        R_vweight = weighted.mean(excess_ret, ME_USD_mil)),
            by = c("Year_ret", "Month", "size_BM")]
    Size_OP_ret <- LHS[, list(R_eweight = mean(excess_ret),
        R_vweight = weighted.mean(excess_ret, ME_USD_mil)),
            by = c("Year_ret", "Month", "size_OP")]
    Size_Inv_ret <- LHS[, list(R_eweight = mean(excess_ret),
        R_vweight = weighted.mean(excess_ret, ME_USD_mil)),
            by = c("Year_ret", "Month", "size_Inv")]

### New Setup
    library(data.table)
    library(plyr)
    RHS <- read.csv("germany_lhs.csv")
    RHS <- data.frame(RHS)

# Determining groups
    RHS <- ddply(RHS,.(Year), transform,
        size_gr = ifelse(size > median(size, na.rm = T), "B", "S"))

    RHS <- ddply(RHS,.(Year), transform,
        BM_gr = ifelse(BM > quantile(BM, 0.7, na.rm = T), "H",
            ifelse(BM < quantile(BM, 0.3, na.rm = T), "L", "N")))
    RHS <- ddply(RHS,.(Year), transform,
        Inv_gr = ifelse(Investment > quantile(Investment, 0.7, na.rm = T),
            "A", ifelse(Investment < quantile(Investment, 0.3, na.rm = T),
                "C", "N")))
    RHS <- ddply(RHS,.(Year), transform,
        OP_gr = ifelse(Profitability > quantile(Profitability, 0.7, na.rm = T),
            "R", ifelse(Profitability < quantile(Profitability, 0.3,
                na.rm = T), "W", "N")))

# Combining groups
    RHS <- within(RHS, size_BM <- paste0(size_gr, BM_gr))
    RHS <- within(RHS, size_OP <- paste0(size_gr, OP_gr))
    RHS <- within(RHS, size_Inv <- paste0(size_gr, Inv_gr))
    RHS <- na.omit(RHS)

# A .csv file is saved here called germany_rhs.csv

### Constructing the SMB factor
```

```
# Computing average return for the size_BM groups
    SMB_BM <- data.table(RHS)
    SMB_BM <- SMB_BM[, list(r_eweight = mean(RETURN), r_vweight =
        weighted.mean(RETURN, ME_USD_mil), ME_USD_mil = sum(ME_USD_mil)),
            by = c("Year_ret", "Month", "size_BM", "size_gr")]

# Computing average return for the size_OP groups
    SMB_OP <- data.table(RHS)
    SMB_OP <- SMB_OP[, list(r_eweight = mean(RETURN), r_vweight =
        weighted.mean(RETURN, ME_USD_mil), ME_USD_mil = sum(ME_USD_mil)),
            by = c("Year_ret", "Month", "size_OP", "size_gr")]

# Computing average return for the size_Inv groups
    SMB_Inv <- data.table(RHS)
    SMB_Inv <- SMB_Inv[, list(r_eweight = mean(RETURN), r_vweight =
        weighted.mean(RETURN, ME_USD_mil), ME_USD_mil = sum(ME_USD_mil)),
            by = c("Year_ret", "Month", "size_Inv", "size_gr")]

# Compute average return by size group
    SMB_BM <- SMB_BM[, list(r_eweight = mean(r_eweight), r_vweight =
        mean(r_vweight)), by = c("Year_ret", "Month", "size_gr")]

# Sort small before big
    SMB_BM <- SMB_BM[order(Year_ret, Month, -size_gr),]

# Compute monthly average return for small minus big (BM)
    SMB_BM <- SMB_BM[, list(SMB_BM_m = diff(-r_vweight)),
        by = c("Year_ret", "Month")]

# Compute average return by size group
    SMB_OP <- SMB_OP[, list(r_eweight = mean(r_eweight), r_vweight =
        mean(r_vweight)), by = c("Year_ret", "Month", "size_gr")]

# Sort small before big
    SMB_OP <- SMB_OP[order(Year_ret, Month, -size_gr),]

# Compute monthly average return for small minus big (OP)
    SMB_OP <- SMB_OP[, list(SMB_OP_m = diff(-r_vweight)),
        by = c("Year_ret", "Month")]

# Compute average return by size group
    SMB_Inv <- SMB_Inv[, list(r_eweight = mean(r_eweight), r_vweight =
        mean(r_vweight)), by = c("Year_ret", "Month", "size_gr")]

# Sort small before big
```

```
    SMB_Inv <- SMB_Inv[order(Year_ret, Month, -size_gr),]

# Compute monthly average return for small minus big (Inv)
    SMB_Inv <- SMB_Inv[, list(SMB_Inv_m = diff(-r_vweight)),
        by = c("Year_ret", "Month")]

# Calculate total SMB based on SMB_BM, SMB_OP, and SMB_Inv
    factor_obs <- data.frame(SMB_BM$Year_ret, SMB_BM$Month, (SMB_BM$SMB_BM_m +
        SMB_OP$SMB_OP_m + SMB_Inv$SMB_Inv_m)/3)
    colnames(factor_obs)[1:3] <- c("Year", "Month", "SMB")

### Constructing the HML factor
    Germany_HML <- read.csv("germany_rhs.csv")
    Germany_HML <- data.table(Germany_HML)
    Germany_HML <- Germany_HML[, list(r_eweight = mean(RETURN), r_vweight =
        weighted.mean(RETURN, ME_USD_mil),
            ME_USD_mil = sum(ME_USD_mil)), by = c("Year_ret", "Month",
                "size_BM", "BM_gr")]

# Compute average return by BM group
    Germany_HML <- Germany_HML[, list(r_eweight = mean(r_eweight), r_vweight =
        mean(r_vweight)), by = c("Year_ret", "Month", "BM_gr")]

# Delete neutral stocks such that there are only the H and L groups left
    Germany_HML <- Germany_HML[!(Germany_HML$BM_gr == "N"),]

# Sort high before low
    Germany_HML <- Germany_HML[order(Year_ret, Month, BM_gr),]

# Compute high minus low
    Germany_HML <- Germany_HML[, list(HML = diff(-r_vweight)),
        by = c("Year_ret", "Month")]

### Constructing the CMA factor
    Germany_CMA <- read.csv("germany_rhs.csv")
    Germany_CMA <- data.table(Germany_CMA)
    Germany_CMA <- Germany_CMA[, list(r_eweight = mean(RETURN), r_vweight =
        weighted.mean(RETURN, ME_USD_mil),
            ME_USD_mil = sum(ME_USD_mil)), by = c("Year_ret", "Month",
                "size_Inv", "Inv_gr")]

# Compute average return by Inv group
    Germany_CMA <- Germany_CMA[, list(r_eweight = mean(r_eweight), r_vweight =
        mean(r_vweight)), by = c("Year_ret", "Month", "Inv_gr")]
```

```
# Delete neutral stocks such that there are only the C and A groups left
    Germany_CMA <- Germany_CMA[!(Germany_CMA$Inv_gr == "N"),]

# Sort such that C is before A
    Germany_CMA <- Germany_CMA[order(Year_ret, Month, -Inv_gr),]

# Compute CMA
    Germany_CMA <- Germany_CMA[, list(CMA = diff(-r_vweight)),
        by = c("Year_ret", "Month")]

### Constructing the RMW factor
    Germany_RMW <- read.csv("germany_rhs.csv")
    Germany_RMW <- data.table(Germany_RMW)
    Germany_RMW <- Germany_RMW[, list(r_eweight = mean(RETURN), r_vweight =
        weighted.mean(RETURN, ME_USD_mil),
            ME_USD_mil = sum(ME_USD_mil)), by = c("Year_ret", "Month",
                "size_OP", "OP_gr")]

# Compute average return by OP group
    Germany_RMW <- Germany_RMW[, list(r_eweight = mean(r_eweight), r_vweight =
        mean(r_vweight)), by = c("Year_ret", "Month", "OP_gr")]

# Delete neutral stocks such that there are only the R and W groups left
    Germany_RMW <- Germany_RMW[!(Germany_RMW$OP_gr == "N"),]

# Sort such that R is before W
    Germany_RMW <- Germany_RMW[order(Year_ret, Month, OP_gr)]

# Compute RMW
    Germany_RMW <- Germany_RMW[, list(RMW = diff(-r_vweight)),
        by = c("Year_ret", "Month")]

### Constructing the EMR factor
    Germany_EMR <- read.csv("Ger_after_trimming_v2.csv")
    Germany_EMR <- data.table(Germany_EMR)
    Germany_EMR <- Germany_EMR[order(Year, Month, company)]
    Germany_EMR <- Germany_EMR[, list(EMR_eweight = mean(excess_ret,
        na.rm = T), EMR_vweight = weighted.mean(excess_ret, ME_USD_mil,
            na.rm = T)), by = c("Year", "Month")]
    Germany_EMR <- Germany_EMR[order(Year, Month)]

### Constructing the WML factor
    mom_a <- read.csv("mom_sorting_variable.csv")
    mom_a <- data.frame(mom_a[rep(seq_len(nrow(mom_a)), each = 12),])
    mom_a$Month <- rep(c(7:12,1:6), times = nrow(mom_a)/12)
```

```
    mom_a <- data.table(mom_a)
    mom_a[, Year_ret := ifelse(Month > 6, year, year + 1)]

#Merge data with return data
    GERMANY_return_data <- read.csv("Ger_after_trimming_v2.csv")
    colnames(GERMANY_return_data)[1] <- "Year_ret"
    mom_a <- merge(mom_a, GERMANY_return_data, by = c("Year_ret", "Month",
        "company"))

#Determine groups
    mom_a <- ddply(mom_a,.(year), transform,
        mom_gr = ifelse(mom > median(mom, na.rm = T), "W", "L"))

    mom_a <- na.omit(mom_a)
    mom_a <- data.table(mom_a)
    mom_a <- mom_a[, list(r_eweight = mean(RETURN), r_vweight =
        weighted.mean(RETURN, ME_USD_mil), ME_USD_mil = sum(ME_USD_mil)),
            by = c("Year_ret", "Month", "mom_gr")]

    mom_a <- mom_a[order(Year_ret, Month, -mom_gr),]
    mom_a <- mom_a[, list(WML = diff(-r_vweight)), by = c("Year_ret", "Month")]
```

RStudio-code written by Christoffer Iversen with inspiration from Hoel and Mix [2016].

Appendix C

R-code: Regression Analysis

```
#Setup
    library(plyr)
    OP_regression <- read.csv("size_OP_merged.csv")

### Three-factor model ###
    OLS_3F <- lapply(9:33, function(x) lm(OP_regression[,x] ~
        OP_regression$EMR + OP_regression$SMB +
            OP_regression$HML, na.action = na.exclude))

    betas_3F <- ldply(OLS_3F, coef)
    colnames(betas_3F)[1:4] <- c("int", "b_EMR", "b_SMB", "b_HML")
    alpha_hat_3F <- as.matrix(betas_3F$int)

    betas_3F <- betas_3F[c(-1)]

    epsilon_hat_3F <- ldply(OLS_3F, residuals)
    sigma_3F <- as.matrix(t(epsilon_hat_3F))
    sigma_3F <- cov(sigma_3F, use = "complete.obs")

#Extracting p-values
    p_val_3F <- lapply(OLS_3F, summary)
    p_val_3F <- lapply(p_val_3F, function(x) x$coef[1,4])
    p_val_3F <- ldply(p_val_3F)
    colnames(p_val_3F)[1] <- "p_int"

#Extracting R_squared
    R_sq_list_3F <- lapply(OLS_3F, summary)
    Adj_R_sq_3F <- matrix(data = NA, nrow = 25, ncol = 1)
        for(i in 1:25){
    Adj_R_sq_3F[i,1] <- as.matrix(R_sq_list_3F[[i]]$adj.r.squared)
```

```
    }

#Calculating GRS test statistic
    T_3F <- as.numeric(length(OP_regression$EMR))
    N_3F <- as.numeric(length(alpha_hat_3F))
    K_3F <- as.numeric(ncol(betas_3F))
    part1_3F <- (T_3F-N_3F-K_3F)/N_3F
    mu_3F <- matrix(nrow = ncol(betas_3F), ncol = 1)
    mu_3F[1,1] <- mean(OP_regression$EMR)
    mu_3F[2,1] <- mean(OP_regression$SMB)
    mu_3F[3,1] <- mean(OP_regression$HML)
    omega_3F <- read.csv("factor_cov_matrix.csv", header = FALSE)
    omega_3F <- omega_3F[c(-4,-5,-6),c(-4,-5,-6)]
    omegaInv_3F <- solve(omega_3F)
    factor_sandwich_3F <- t(mu_3F) %*% solve(omega_3F) %*% mu_3F
    part2_3F <- (1 + as.numeric(factor_sandwich_3F))^-1
    part3_3F <- as.numeric(t(alpha_hat_3F) %*% solve(sigma_3F)
        %*% alpha_hat_3F)

    GRS_3F <- part1_3F*part2_3F*part3_3F
    GRS_pval_3F <- 1-pf(GRS_3F,N_3F,(T_3F-N_3F-K_3F))
    avg_abs_alpha_3F <- mean(abs(alpha_hat_3F))

    alternatively <- (T_3F/N_3F)*((T_3F-N_3F-K_3F)/(T_3F-K_3F-1))
    GRS_alternatively_3F <- alternatively*part2_3F*part3_3F

### Four-factor model ###
    OLS_4F <- lapply(9:33, function(x) lm(OP_regression[,x] ~
        OP_regression$EMR + OP_regression$SMB + OP_regression$HML +
            OP_regression$WML, na.action = na.exclude))

    betas_4F <- ldply(OLS_4F, coef)
    colnames(betas_4F)[1:4] <- c("int", "b_EMR", "b_SMB", "b_HML")
    alpha_hat_4F <- as.matrix(betas_4F$int)

    betas_4F <- betas_4F[c(-1)]

    epsilon_hat_4F <- ldply(OLS_4F, residuals)
    sigma_4F <- as.matrix(t(epsilon_hat_4F))
    sigma_4F <- cov(sigma_4F, use = "complete.obs")

#Extracting p-values
    p_val_4F <- lapply(OLS_4F, summary)
    p_val_4F <- lapply(p_val_4F, function(x) x$coef[1,4])
    p_val_4F <- ldply(p_val_4F)
```

```
    colnames(p_val_4F)[1] <- "p_int"

#Extracting R_squared
    R_sq_list_4F <- lapply(OLS_4F, summary)
    Adj_R_sq_4F <- matrix(data = NA, nrow = 25, ncol = 1)
    for(i in 1:25){
      Adj_R_sq_4F[i,1] <- as.matrix(R_sq_list_4F[[i]]$adj.r.squared)
    }

#Calculating GRS test statistic
    T_4F <- as.numeric(length(OP_regression$EMR))
    N_4F <- as.numeric(length(alpha_hat_4F))
    K_4F <- as.numeric(ncol(betas_4F))
    part1_4F <- (T_4F-N_4F-K_4F)/N_4F
    mu_4F <- matrix(nrow = ncol(betas_4F), ncol = 1)
    mu_4F[1,1] <- mean(OP_regression$EMR)
    mu_4F[2,1] <- mean(OP_regression$SMB)
    mu_4F[3,1] <- mean(OP_regression$HML)
    mu_4F[4,1] <- mean(OP_regression$WML)
    omega_4F <- read.csv("factor_cov_matrix.csv", header = FALSE)
    omega_4F <- omega_4F[c(-4,-5),c(-4,-5)]
    omegaInv_4F <- solve(omega_4F)
    factor_sandwich_4F <- t(mu_4F) %*% solve(omega_4F) %*% mu_4F
    part2_4F <- (1 + as.numeric(factor_sandwich_4F))^-1
    part3_4F <- as.numeric(t(alpha_hat_4F) %*% solve(sigma_4F)
        %*% alpha_hat_4F)

    GRS_4F <- part1_4F*part2_4F*part3_4F
    GRS_pval_4F <- 1-pf(GRS_4F,N_4F,(T_4F-N_4F-K_4F))
    avg_abs_alpha_4F <- mean(abs(alpha_hat_4F))

    alternatively <- (T_4F/N_4F)*((T_4F-N_4F-K_4F)/(T_4F-K_4F-1))
    GRS_alternatively_4F <- alternatively*part2_4F*part3_4F

### Five-factor model ###
    OLS_5F <- lapply(9:33, function(x) lm(OP_regression[,x] ~
        OP_regression$EMR + OP_regression$SMB + OP_regression$HML +
            OP_regression$RMW + OP_regression$CMA,
                na.action = na.exclude))

    betas_5F <- ldply(OLS_5F, coef)
    colnames(betas_5F)[1:4] <- c("int", "b_EMR", "b_SMB", "b_HML")
    alpha_hat_5F <- as.matrix(betas_5F$int)

    betas_5F <- betas_5F[c(-1)]
```

```
    epsilon_hat_5F <- ldply(OLS_5F, residuals)
    sigma_5F <- as.matrix(t(epsilon_hat_5F))
    sigma_5F <- cov(sigma_5F, use = "complete.obs")

#Extracting p-values
    p_val_5F <- lapply(OLS_5F, summary)
    p_val_5F <- lapply(p_val_5F, function(x) x$coef[1,4])
    p_val_5F <- ldply(p_val_5F)
    colnames(p_val_5F)[1] <- "p_int"

#Extracting R_squared
    R_sq_list_5F <- lapply(OLS_5F, summary)
    Adj_R_sq_5F <- matrix(data = NA, nrow = 25, ncol = 1)
    for(i in 1:25){
      Adj_R_sq_5F[i,1] <- as.matrix(R_sq_list_5F[[i]]$adj.r.squared)
    }

#Calculating GRS test statistic
    T_5F <- as.numeric(length(OP_regression$EMR))
    N_5F <- as.numeric(length(alpha_hat_5F))
    K_5F <- as.numeric(ncol(betas_5F))
    part1_5F <- (T_5F-N_5F-K_5F)/N_5F
    mu_5F <- matrix(nrow = ncol(betas_5F), ncol = 1)
    mu_5F[1,1] <- mean(OP_regression$EMR)
    mu_5F[2,1] <- mean(OP_regression$SMB)
    mu_5F[3,1] <- mean(OP_regression$HML)
    mu_5F[4,1] <- mean(OP_regression$RMW)
    mu_5F[5,1] <- mean(OP_regression$CMA)
    omega_5F <- read.csv("factor_cov_matrix.csv", header = FALSE)
    omega_5F <- omega_5F[c(-6),c(-6)]
    omegaInv_5F <- solve(omega_5F)
    factor_sandwich_5F <- t(mu_5F) %*% solve(omega_5F) %*% mu_5F
    part2_5F <- (1 + as.numeric(factor_sandwich_5F))^-1
    part3_5F <- as.numeric(t(alpha_hat_5F) %*% solve(sigma_5F)
        %*% alpha_hat_5F)

    GRS_5F <- part1_5F*part2_5F*part3_5F
    GRS_pval_5F <- 1-pf(GRS_5F,N_5F,(T_5F-N_5F-K_5F))
    avg_abs_alpha_5F <- mean(abs(alpha_hat_5F))

    alternatively <- (T_5F/N_5F)*((T_5F-N_5F-K_5F)/(T_5F-K_5F-1))
    GRS_alternatively_5F <- alternatively*part2_5F*part3_5F

### Six-factor model ###
```

```
    OLS_6F <- lapply(9:33, function(x) lm(OP_regression[,x] ~
        OP_regression$EMR + OP_regression$SMB + OP_regression$HML +
            OP_regression$RMW + OP_regression$CMA + OP_regression$WML,
                na.action = na.exclude))

    betas_6F <- ldply(OLS_6F, coef)
    colnames(betas_6F)[1:4] <- c("int", "b_EMR", "b_SMB", "b_HML")
    alpha_hat_6F <- as.matrix(betas_6F$int)

    betas_6F <- betas_6F[c(-1)]

    epsilon_hat_6F <- ldply(OLS_6F, residuals)
    sigma_6F <- as.matrix(t(epsilon_hat_6F))
    sigma_6F <- cov(sigma_6F, use = "complete.obs")

#Extracting p-values
    p_val_6F <- lapply(OLS_6F, summary)
    p_val_6F <- lapply(p_val_6F, function(x) x$coef[1,4])
    p_val_6F <- ldply(p_val_6F)
    colnames(p_val_6F)[1] <- "p_int"

#Extracting R_squared
    R_sq_list_6F <- lapply(OLS_6F, summary)
    Adj_R_sq_6F <- matrix(data = NA, nrow = 25, ncol = 1)
    for(i in 1:25){
      Adj_R_sq_6F[i,1] <- as.matrix(R_sq_list_6F[[i]]$adj.r.squared)
    }

#Calculating GRS test statistic
    T_6F <- as.numeric(length(OP_regression$EMR))
    N_6F <- as.numeric(length(alpha_hat_6F))
    K_6F <- as.numeric(ncol(betas_6F))
    part1_6F <- (T_6F-N_6F-K_6F)/N_6F
    mu_6F <- matrix(nrow = ncol(betas_6F), ncol = 1)
    mu_6F[1,1] <- mean(OP_regression$EMR)
    mu_6F[2,1] <- mean(OP_regression$SMB)
    mu_6F[3,1] <- mean(OP_regression$HML)
    mu_6F[4,1] <- mean(OP_regression$RMW)
    mu_6F[5,1] <- mean(OP_regression$CMA)
    mu_6F[6,1] <- mean(OP_regression$WML)
    omega_6F <- read.csv("factor_cov_matrix.csv", header = FALSE)
    omegaInv_6F <- solve(omega_6F)
    factor_sandwich_6F <- t(mu_6F) %*% solve(omega_6F) %*% mu_6F
    part2_6F <- (1 + as.numeric(factor_sandwich_6F))^-1
    part3_6F <- as.numeric(t(alpha_hat_6F) %*% solve(sigma_6F)
```

```
        %*% alpha_hat_6F)

    GRS_6F <- part1_6F*part2_6F*part3_6F
    GRS_pval_6F <- 1-pf(GRS_6F,N_6F,(T_6F-N_6F-K_6F))
    avg_abs_alpha_6F <- mean(abs(alpha_hat_6F))

    alternatively <- (T_6F/N_6F)*((T_6F-N_6F-K_6F)/(T_6F-K_6F-1))
    GRS_alternatively_6F <- alternatively*part2_6F*part3_6F

### To save space, only the programming code for size and operating
    profitability is shown here. The same thing is repeated for
    size and book-to-market sorts, size and investment sorts, and
    all 75 portfolios together ###


### New Setup for cross-sectional regressions
    library(plyr)
    library(MASS)

# Five-factor model
    CS_regression <- read.csv("CS_all_portfolios_5F.csv")
    CS_5F <- lapply(6:323, function(x) lm(CS_regression[,x] ~
        CS_regression$b_EMR_5F + CS_regression$b_SMB_5F +
            CS_regression$b_HML_5F + CS_regression$b_RMW_5F +
                CS_regression$b_CMA_5F + 0, na.action = na.exclude))

    lambdas_5F <- ldply(CS_5F, coef)
    colnames(lambdas_5F)[1:5] <- c("l_EMR",
        "l_SMB", "l_HML", "l_RMW", "l_CMA")
    l_EMR_5F <- mean(lambdas_5F$l_EMR)
    l_SMB_5F <- mean(lambdas_5F$l_SMB)
    l_HML_5F <- mean(lambdas_5F$l_HML)
    l_RMW_5F <- mean(lambdas_5F$l_RMW)
    l_CMA_5F <- mean(lambdas_5F$l_CMA)

    t_squared <- length(lambdas_5F$l_EMR)^2

    lambdas_5F$sqdev_EMR_5F <- (lambdas_5F$l_EMR - l_EMR_5F)^2
    se_EMR_5F <-  sqrt(1/t_squared * sum(lambdas_5F$sqdev_EMR_5F))
    t_stat_EMR_5F <- l_EMR_5F/se_EMR_5F

    lambdas_5F$sqdev_SMB_5F <- (lambdas_5F$l_SMB - l_SMB_5F)^2
    se_SMB_5F <-  sqrt(1/t_squared * sum(lambdas_5F$sqdev_SMB_5F))
    t_stat_SMB_5F <- l_SMB_5F/se_SMB_5F
```

```
    lambdas_5F$sqdev_HML_5F <- (lambdas_5F$l_HML - l_HML_5F)^2
    se_HML_5F <-  sqrt(1/t_squared * sum(lambdas_5F$sqdev_HML_5F))
    t_stat_HML_5F <- l_HML_5F/se_HML_5F

    lambdas_5F$sqdev_RMW_5F <- (lambdas_5F$l_RMW - l_RMW_5F)^2
    se_RMW_5F <-  sqrt(1/t_squared * sum(lambdas_5F$sqdev_RMW_5F))
    t_stat_RMW_5F <- l_RMW_5F/se_RMW_5F

    lambdas_5F$sqdev_CMA_5F <- (lambdas_5F$l_CMA - l_CMA_5F)^2
    se_CMA_5F <-  sqrt(1/t_squared * sum(lambdas_5F$sqdev_CMA_5F))
    t_stat_CMA_5F <- l_CMA_5F/se_CMA_5F

    residuals_alpha_5F <- as.matrix(ldply(CS_5F, residuals))
    cochrane_alpha <- as.matrix(colMeans(residuals_alpha_5F, na.rm = T))
    cochrane_alpha_t <- t(as.matrix(ldply(CS_5F, residuals)))
    cochrane_alpha_t[is.na(cochrane_alpha_t)] <- 0
    cochrane_alpha_filled <- matrix(cochrane_alpha, nrow =
        length(cochrane_alpha), ncol = 318, byrow = FALSE)
    cochrane_1 <- cochrane_alpha_t - cochrane_alpha_filled
    cochrane_2 <- t(cochrane_1)
    cochrane_3 <- (cochrane_1 %*% cochrane_2)/t_squared
    cochrane_4 <- ginv(cochrane_3, tol = 0.1)
    cochrane_5 <- t(cochrane_alpha) %*% cochrane_4 %*% cochrane_alpha

# Six-factor model
    CS_regression <- read.csv("CS_all_portfolios_6F.csv")
    CS_6F <- lapply(7:324, function(x) lm(CS_regression[,x] ~
        0 + CS_regression$b_EMR_6F + CS_regression$b_SMB_6F +
            CS_regression$b_HML_6F + CS_regression$b_RMW_6F +
                CS_regression$b_CMA_6F + CS_regression$b_WML_6F,
                    na.action = na.exclude))

    lambdas_6F <- ldply(CS_6F, coef)
    colnames(lambdas_6F)[1:6] <- c("l_EMR", "l_SMB",
        "l_HML", "l_RMW", "l_CMA", "l_WML")

    l_EMR_6F <- mean(lambdas_6F$l_EMR)
    l_SMB_6F <- mean(lambdas_6F$l_SMB)
    l_HML_6F <- mean(lambdas_6F$l_HML)
    l_RMW_6F <- mean(lambdas_6F$l_RMW)
    l_CMA_6F <- mean(lambdas_6F$l_CMA)
    l_WML_6F <- mean(lambdas_6F$l_WML)

    t_squared <- length(lambdas_6F$l_EMR)^2
```

```
    lambdas_6F$sqdev_EMR_6F <- (lambdas_6F$l_EMR - l_EMR_6F)^2
    se_EMR_6F <-  sqrt(1/t_squared * sum(lambdas_6F$sqdev_EMR_6F))
    t_stat_EMR_6F <- l_EMR_6F/se_EMR_6F

    lambdas_6F$sqdev_SMB_6F <- (lambdas_6F$l_SMB - l_SMB_6F)^2
    se_SMB_6F <-  sqrt(1/t_squared * sum(lambdas_6F$sqdev_SMB_6F))
    t_stat_SMB_6F <- l_SMB_6F/se_SMB_6F

    lambdas_6F$sqdev_HML_6F <- (lambdas_6F$l_HML - l_HML_6F)^2
    se_HML_6F <-  sqrt(1/t_squared * sum(lambdas_6F$sqdev_HML_6F))
    t_stat_HML_6F <- l_HML_6F/se_HML_6F

    lambdas_6F$sqdev_RMW_6F <- (lambdas_6F$l_RMW - l_RMW_6F)^2
    se_RMW_6F <-  sqrt(1/t_squared * sum(lambdas_6F$sqdev_RMW_6F))
    t_stat_RMW_6F <- l_RMW_6F/se_RMW_6F

    lambdas_6F$sqdev_CMA_6F <- (lambdas_6F$l_CMA - l_CMA_6F)^2
    se_CMA_6F <-  sqrt(1/t_squared * sum(lambdas_6F$sqdev_CMA_6F))
    t_stat_CMA_6F <- l_CMA_6F/se_CMA_6F

    lambdas_6F$sqdev_WML_6F <- (lambdas_6F$l_WML - l_WML_6F)^2
    se_WML_6F <- sqrt(1/t_squared * sum(lambdas_6F$sqdev_WML_6F))
    t_stat_WML_6F <- l_WML_6F/se_WML_6F

    residuals_alpha_6F <- as.matrix(ldply(CS_6F, residuals))
    cochrane_alpha <- as.matrix(colMeans(residuals_alpha_6F, na.rm = T))
    cochrane_alpha_t <- t(as.matrix(ldply(CS_6F, residuals)))
    cochrane_alpha_t[is.na(cochrane_alpha_t)] <- 0
    cochrane_alpha_filled <- matrix(cochrane_alpha, nrow =
        length(cochrane_alpha), ncol = 318, byrow = FALSE)
    cochrane_1 <- cochrane_alpha_t - cochrane_alpha_filled
    cochrane_2 <- t(cochrane_1)
    cochrane_3 <- (cochrane_1 %*% cochrane_2)/t_squared
    cochrane_4 <- ginv(cochrane_3, tol = (0.1))
    cochrane_5 <- t(cochrane_alpha) %*% cochrane_4 %*% cochrane_alpha

# Three-factor model
    CS_regression <- read.csv("CS_all_portfolios_3F.csv")

    CS_3F <- lapply(4:321, function(x) lm(CS_regression[,x] ~
        0 + CS_regression$b_EMR_3F + CS_regression$b_SMB_3F +
            CS_regression$b_HML_3F, na.action = na.exclude))

    lambdas_3F <- ldply(CS_3F, coef)
    colnames(lambdas_3F)[1:3] <- c("l_EMR", "l_SMB", "l_HML")
```

```
    l_EMR_3F <- mean(lambdas_3F$l_EMR)
    l_SMB_3F <- mean(lambdas_3F$l_SMB)
    l_HML_3F <- mean(lambdas_3F$l_HML)

    t_squared <- length(lambdas_3F$l_EMR)^2

    lambdas_3F$sqdev_EMR_3F <- (lambdas_3F$l_EMR - l_EMR_3F)^2
    se_EMR_3F <-  sqrt(1/t_squared * sum(lambdas_3F$sqdev_EMR_3F))
    t_stat_EMR_3F <- l_EMR_3F/se_EMR_3F

    lambdas_3F$sqdev_SMB_3F <- (lambdas_3F$l_SMB - l_SMB_3F)^2
    se_SMB_3F <-  sqrt(1/t_squared * sum(lambdas_3F$sqdev_SMB_3F))
    t_stat_SMB_3F <- l_SMB_3F/se_SMB_3F

    lambdas_3F$sqdev_HML_3F <- (lambdas_3F$l_HML - l_HML_3F)^2
    se_HML_3F <-  sqrt(1/t_squared * sum(lambdas_3F$sqdev_HML_3F))
    t_stat_HML_3F <- l_HML_3F/se_HML_3F

    residuals_alpha_3F <- as.matrix(ldply(CS_3F, residuals))
    cochrane_alpha <- as.matrix(colMeans(residuals_alpha_3F, na.rm = T))
    cochrane_alpha_t <- t(as.matrix(ldply(CS_3F, residuals)))
    cochrane_alpha_t[is.na(cochrane_alpha_t)] <- 0
    cochrane_alpha_filled <- matrix(cochrane_alpha, nrow =
        length(cochrane_alpha), ncol = 318, byrow = FALSE)
    cochrane_1 <- cochrane_alpha_t - cochrane_alpha_filled
    cochrane_2 <- t(cochrane_1)
    cochrane_3 <- (cochrane_1 %*% cochrane_2)/t_squared
    cochrane_4 <- ginv(cochrane_3, tol = (0.1))
    cochrane_5 <- t(cochrane_alpha) %*% cochrane_4 %*% cochrane_alpha

# Four-factor model
    CS_regression <- read.csv("CS_all_portfolios_4F.csv")

    CS_4F <- lapply(5:322, function(x) lm(CS_regression[,x] ~
        0 + CS_regression$b_EMR_4F + CS_regression$b_SMB_4F +
            CS_regression$b_HML_4F + CS_regression$b_WML_4F,
                na.action = na.exclude))

    lambdas_4F <- ldply(CS_4F, coef)
    colnames(lambdas_4F)[1:4] <- c("l_EMR", "l_SMB", "l_HML", "l_WML")

    l_EMR_4F <- mean(lambdas_4F$l_EMR)
    l_SMB_4F <- mean(lambdas_4F$l_SMB)
    l_HML_4F <- mean(lambdas_4F$l_HML)
```

```
l_WML_4F <- mean(lambdas_4F$l_WML)

t_squared <- length(lambdas_4F$l_EMR)^2

lambdas_4F$sqdev_EMR_4F <- (lambdas_4F$l_EMR - l_EMR_4F)^2
se_EMR_4F <-  sqrt(1/t_squared * sum(lambdas_4F$sqdev_EMR_4F))
t_stat_EMR_4F <- l_EMR_4F/se_EMR_4F
lambdas_4F$sqdev_SMB_4F <- (lambdas_4F$l_SMB - l_SMB_4F)^2
se_SMB_4F <-  sqrt(1/t_squared * sum(lambdas_4F$sqdev_SMB_4F))
t_stat_SMB_4F <- l_SMB_4F/se_SMB_4F
lambdas_4F$sqdev_HML_4F <- (lambdas_4F$l_HML - l_HML_4F)^2
se_HML_4F <-  sqrt(1/t_squared * sum(lambdas_4F$sqdev_HML_4F))
t_stat_HML_4F <- l_HML_4F/se_HML_4F
lambdas_4F$sqdev_WML_4F <- (lambdas_4F$l_WML - l_WML_4F)^2
se_WML_4F <- sqrt(1/t_squared * sum(lambdas_4F$sqdev_WML_4F))
t_stat_WML_4F <- l_WML_4F/se_WML_4F

residuals_alpha_4F <- as.matrix(ldply(CS_4F, residuals))
cochrane_alpha <- as.matrix(colMeans(residuals_alpha_4F, na.rm = T))
cochrane_alpha_t <- t(as.matrix(ldply(CS_4F, residuals)))
cochrane_alpha_t[is.na(cochrane_alpha_t)] <- 0
cochrane_alpha_filled <- matrix(cochrane_alpha, nrow =
    length(cochrane_alpha), ncol = 318, byrow = FALSE)
cochrane_1 <- cochrane_alpha_t - cochrane_alpha_filled
cochrane_2 <- t(cochrane_1)
cochrane_3 <- (cochrane_1 %*% cochrane_2)/t_squared
cochrane_4 <- ginv(cochrane_3, tol = (0.1))
cochrane_5 <- t(cochrane_alpha) %*% cochrane_4 %*% cochrane_alpha
```

RStudio-code written by Christoffer Iversen.

Bibliography

Gil Aharoni, Bruce Grundy, and Qi Zeng. Stock returns and the miller modigliani valuation formula: Revisiting the fama french analysis. *Journal of Financial Economics*, 110(2):347–357, 2013.

Andrew Ang, Jun Liu, and Krista Schwarz. Using individual stocks or portfolios in tests of factor models. *Working Paper*, 2017.

Sabine Artmann, Philipp Finter, Alexander Kempf, Stefan Koch, and Erik Theissen. The cross-section of german stock returns: New data and new evidence. *Schmalenbach Business Review*, 64(1):20–43, 2012.

Doron Avramov, Tarun Chordia, Gergana Jostova, and Alexander Philipov. Momentum and credit rating. *The Journal of Finance*, 62(5):2503–2520, 2007.

Rolf W Banz. The relationship between return and market value of common stocks. *Journal of financial economics*, 9(1):3–18, 1981.

David Blitz. *Benchmarking Benchmarks*. PhD thesis, Erasmus University Rotterdam, 3 2011.

David Blitz, Matthias X. Hanauer, Milan Vidojevic, and Pim van Vliet. Five concerns with the five-factor model. *Working Paper*, 2017.

Marshall E Blume. Portfolio theory: a step toward its practical application. *The Journal of Business*, 43(2):152–173, 1970.

Werner FM Bondt and Richard Thaler. Does the stock market overreact? *The Journal of finance*, 40(3):793–805, 1985.

Alan Bryman. *Social Research Methods 4th edition*. Oxford University Press, 2012.

Roman Brückner, Patrick Lehmann, Martin H. Schmidt, and Richard Stehle. Non-u.s. multi-factor data sets should be used with caution. *Working Paper*, 2015.

John Y Campbell. Asset pricing at the millennium. *The Journal of Finance*, 55 (4):1515–1567, 2000.

Mark M Carhart. On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82, 1997.

Murray Carlson, Adlai Fisher, and Ron Giammarino. Corporate investment and asset price dynamics: implications for the cross-section of returns. *The Journal of Finance*, 59(6):2577–2603, 2004.

KC Chan and Nai-Fu Chen. Structural and return characteristics of small and large firms. *The Journal of Finance*, 46(4):1467–1484, 1991.

Louis KC Chan, Narasimhan Jegadeesh, and Josef Lakonishok. Momentum strategies. *The Journal of Finance*, 51(5):1681–1713, 1996.

John H. Cochrane. *Asset Pricing Revised Edition*. Princeton University Press, 2005.

Ilan Cooper and Richard Priestley. Real investment and risk dynamics. *Journal of Financial Economics*, 101(1):182–205, 2011.

Eugene F Fama and Kenneth R French. Permanent and temporary components of stock prices. *Journal of political Economy*, 96(2):246–273, 1988.

Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.

Eugene F Fama and Kenneth R French. Size and book-to-market factors in earnings and returns. *The journal of finance*, 50(1):131–155, 1995.

Eugene F Fama and Kenneth R French. Multifactor explanations of asset pricing anomalies. *The journal of finance*, 51(1):55–84, 1996.

Eugene F Fama and Kenneth R French. Value versus growth: The international evidence. *The journal of finance*, 53(6):1975–1999, 1998.

Eugene F Fama and Kenneth R French. Profitability, investment and average returns. *Journal of Financial Economics*, 82(3):491–518, 2006.

Eugene F Fama and Kenneth R French. Size, value, and momentum in international stock returns. *Journal of financial economics*, 105(3):457–472, 2012.

Eugene F. Fama and Kenneth R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015.

Eugene F Fama and Kenneth R French. International tests of a five-factor asset pricing model. *Journal of financial Economics*, 123(3):441–463, 2017.

Eugene F Fama and James D MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3):607–636, 1973.

James Foye. Testing alternative versions of the fama-french five-factor model in the uk. *Working Paper*, 2017.

James Foye, Dušan Mramor, and Marko Pahor. A respecified fama french three-factor model for the new european union member states. *Journal of International Financial Management & Accounting*, 24(1):3–25, 2013.

Kenneth R. French. Data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, 2018. Online, accessed 31-01-2018.

Michael R Gibbons, Stephen A Ross, and Jay Shanken. A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152, 1989.

Myron J Gordon. Dividends, earnings, and stock prices. *The review of economics and statistics*, pages 99–105, 1959.

Amit Goyal. Empirical cross-sectional asset pricing: a survey. *Financial Markets and Portfolio Management*, 26(1):3–38, 2012.

John M Griffin, Xiuqing Ji, and J Spencer Martin. Momentum investing and business cycle risk: Evidence from pole to pole. *The Journal of Finance*, 58(6): 2515–2547, 2003.

Campbell R Harvey, Yan Liu, and Heqing Zhu. … and the cross-section of expected returns. *The Review of Financial Studies*, 29(1):5–68, 2016.

Werner Heisenberg. Positivism, metaphysics and religion. *Physics and Beyond: Encounters and Conversations*, pages 205–217, 1971.

Alexander R. Hoel and Fabia R. Mix. How suitable is the fama-french five-factor model for describing german and norwegian stock returns? Master's thesis, Norwegian School of Economics, 2016.

Harrison Hong, Terence Lim, and Jeremy C Stein. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance*, 55(1):265–295, 2000.

Soeren Hvidkjaer. A trade-based analysis of momentum. *The Review of Financial Studies*, 19(2):457–491, 2006.

Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91, 1993.

Narasimhan Jegadeesh and Sheridan Titman. Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of finance*, 56(2): 699–720, 2001.

Michael C Jensen, Fischer Black, and Myron S Scholes. The capital asset pricing model: Some empirical tests. *Studies in the Theory of Capital Markets*, 1972.

Roger W Klein and Vijay S Bawa. The effect of limited information and estimation risk on optimal portfolio diversification. *Journal of Financial Economics*, 5(1): 89–111, 1977.

Charles Lee and Bhaskaran Swaminathan. Price momentum and trading volume. *the Journal of Finance*, 55(5):2017–2069, 2000.

Jimmy Liew and Maria Vassalou. Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics*, 57 (2):221–245, 2000.

Thomas R Lindlof and Bryan C Taylor. *Qualitative communication research methods*. Sage publications, 2017.

John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, pages 13–37, 1965.

Merton H Miller and Franco Modigliani. Dividend policy, growth, and the valuation of shares. *the Journal of Business*, 34(4):411–433, 1961.

Jan Mossin. Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, pages 768–783, 1966.

Claus Munk. Financial markets and investments (lecture notes in the course "investments" at copenhagen business school), 2016.

Robert Novy-Marx. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28, 2013.

Einar L. Olsen. Factor effects in nordic equities. Master's thesis, Copenhagen Business School, 2016.

Barr Rosenberg, Kenneth Reid, and Ronald Lanstein. Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3):9–16, 1985.

Stephen A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360, December 1976.

Bruce Russell. A priori justification and knowledge. *Stanford Encyclopedia of Philosophy*, 2010.

G William Schwert. Anomalies and market efficiency. *Handbook of the Economics of Finance*, 1:939–974, 2003.

William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.

Robert Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3):421–36, 1981.

Robert Shiller. *Market volatility*. MIT press, 1992.

Dennis Stattman. Book values and stock returns. *The Chicago MBA: A journal of selected papers*, 4(1):25–45, 1980.

Lei Sun, Kuo-Chiang John Wei, and Feixue Xie. On the explanations for the gross profitability effect: Insights from international equity markets. *AsianFA Conference Paper*, 2014.

Sheridan Titman, KC John Wei, and Feixue Xie. Capital investments and stock returns. *Journal of financial and Quantitative Analysis*, 39(4):677–700, 2004.

Huijun Wang and Jianfeng Yu. Dissecting the profitability premium. *Working Paper*, 2013.

Lu Zhang. The value premium. *The Journal of Finance*, 60(1):67–103, 2005.