BI Norwegian Business School - campus Oslo

GRA 19502

Master Thesis

Component of continuous assessment: Forprosjekt, Thesis MSc

The Five-Factor Asset Pricing Model: A Corporate Finance point of view

Navn:	Dino Priganica, Ikram Ul-Haq Rashid

Start:	01.01.2018 09.00
Finish:	15.01.2018 12.00

Title: *The Five-Factor Asset Pricing Model: A Corporate Finance point of view*

Study programme: *Msc in Business, Specialization in Finance*

Supervisor: *Paul Ehling*

Table of Contents

Abstract

This research paper looks at the work of Eugene F. Fama and Kenneth R. French on asset pricing models and sets out to tests the relationship between average asset returns over time against various risk factors. It applies the well-known and well-established Gordon Growth Valuation model from corporate finance to uncover risk factors which can potentially have a statistically significant effect on asset returns, after controlling for the factors previously identified by Fama and French (2015). In a sense, the paper is set to (i) confirm the validity of the Fama and French (2015) Five-Factor model and (ii) identify additional risk factors with statistical significance. To keep in line with the works of Fama and French (1993, 1994, 1996, 2008 and 2015), we will apply the methodology outlined by Fama and MacBeth (1973) and run a Fama-MacBeth cross-sectional analysis of stocks listed on The New York Stock Exchange (''NYSE'').

1.0 Introduction to the research topic

1.1 Previous theory on the subject

Academics and practitioners have forever tried to create a model for the pricing of securities. The father of all such models is the Capital Asset Pricing Model (''CAPM''), developed in parallel by Sharpe (1964), Lintner (1965) and Black (1972), hence why it is often referred to as the Sharpe-Lintner-Black Model (''SLB''). According to Fischer Black (1972), the SLB-model states that any capital asset for a single period, and given certain assumptions, will satisfy the following equation:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f]$$

Where $E(R_i)$ is defined as the return on asset *i* for the period; R_f is the return of a risk-free asset for the period; $E(R_m)$ is the return of the market portfolio (all assets taken together); β_i is the market sensitivity of asset *i* (Black, 1972). By further expanding the equation we observe the following:

$$E(R_i) = \frac{(S_1 + Dividends + Interest + other\ contributions) - S_0}{S_0}$$

$$\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)}$$

Hence, the basic intuition behind the SLB-model is that the only factor that contributes to asset returns over and above the risk-free rate is the exposure to systematic risk. I.e., the *more correlated* asset *i* is with the market portfolio (e.g. market index), the higher the systematic risk in which you are rewarded with a higher expected return on the asset.

Even though the classical SLB-model is highly regarded amongst finance professionals, many researchers have argued that there are some patterns in average stock returns that cannot be explained using this model (Fama & French, 1996). Such patterns are referred to as anomalies. Here are some examples of such anomalies.

- DeBondt and Thaler (1985): Reversal in long-term returns
- Jegadeesh and Titman (1993): Short-term returns tend to continue
- Banz (1981): Average stock return is related to firm size
- Basu (1983): Book-to-market-equity

To capture such anomalies and model them, Fama and French (1996) created the now famous three-factor model. Mathematically, the model can be expressed as

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] + s_i E(SMB) + h_i E(HML)$$

and is, as such, a direct extension of the older SLB-model. Fama and French (1993) argued that the factors contained within this model will capture many of the CAPM average-return anomalies because these anomalies are highly related. Intuitively, the model tries to explain the expected return of asset *i* as a result of the interplay between three factors (Fama and French, 1996):

I. The excess return on a broad market portfolio $(R_m - R_f)$
II. The difference between the return on a portfolio of small stocks and the return on a portfolio of large stocks (SMB, small minus big)
III. The difference between the return on a portfolio of high-book-to-market stocks and the return on a portfolio of low-book-to-market stocks (HML, high minus low)

Fama and French, through their several papers, have indeed proven that many of the anomalies are modelled through the three-factor equation (Fama & French, 1996). Fama and French (1993) find evidence that the model is a good description of returns on portfolios formed on size and BE/ME. Fama and French (1994) use the model to explain industry returns. Fama and French (1996) use the model to show that it captures the returns to portfolios formed on earnings/price, cash flow/price and sales growth.

In the recent years, researchers have found evidence that the Three-factor model is an incomplete representation of expected return because it does not consider returns related to both profitability and investment (See Novy-Marx (2013) and Titman, Wei & Xie (2004)). Fama and French (2015) sets to incorporate these factors by adding two additional factors to the Three-factor model and hence introducing a Five-Factor model. They represent the model as following.

$$R_{i,t} - R_{F,t} = a_i + b_i(R_{M,t} - R_{F,t}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{i,t}$$

We observe two new explanatory variables; RMW, which is the difference between the returns on diversified portfolios of stocks with robust and weak profitability; CMA, which is the difference between the returns on diversified portfolios of the stocks of low and high investment firms (Fama & French, 2015).

1.2 Our thesis

This paper is set to apply research by Sharpe (1964), Lintner (1965), Black (1972), Fama and French (1993, 1994, 1996, 2008 and 2015) in an attempt to uncover and model additional variables which might potentially add to the explanatory power of such models. I.e. we set to study additional factors that potentially influence the return of individual securities, after controlling for the variables identified by previous research.

The SML-model and the Five-factor model are essentially *asset pricing models*, i.e. many practitioners within corporate finance apply such models to calculate the discount rate and apply these to various valuation-models to finally calculate the value of the company. The most widely known model of this sort is the Gordon Growth Model (Koller & Goedhart, 2015). The Gordon Growth Model (''GGM'') can be expressed mathematically as;

$$TEV = \frac{NOPLAT_{t+1}(1 - {Growth}/{ROIC})}{WACC - Growth}$$

Where TEV is the terminal enterprise value of the company; NOPLAT is the net operating profits less adjusted taxes; WACC is the weighted average cost of capital; ROIC is return on invested capital; growth is an expression of revenue growth.

These variables can themselves be expressed mathematically in the following way (Koller & Goedhart, 2015);

- $NOPLAT = EBIT * (1 - T_c)$
- $WACC = \left(\frac{E}{V}\right) * R_e + \left(\frac{D}{V}\right) * R_d * (1 - T_c)$
- ROIC = $\frac{EBIT*(1-T_c)}{Invested\ Capital}$
- $Sales\ growth = \frac{Revenue_t - Revenue_{t-1}}{Revenue_{t-1}}$

Hence, to identify variables which can influence the return of individual securities, and thus also the value of the company, we will apply the use of well established models from corporate finance. The basic intuition behind this is that the terminal value of the company is essentially a component of the expected return, hence variables contained within the formula of the terminal value should ultimately influence the expected return. Fama and French (2015) applied the same methodology to identify the variables incorporated in the Five-Factor Model, although they used the Dividend Discount Model. We illustrate this mathematically below.

$$E(R_i) = \frac{(S_1 + Dividends + Interest + other\ contributions) - S_0}{S_0}$$

Assuming no dividends, interest or other contributions

$$E(R_i) = \frac{S_1 - S_0}{S_0}$$

Recalling the formula for stock price

$$S_i = \frac{Equity}{Number\ of\ shares\ outstanding}$$

$$S_i = \frac{Terminal\ Value - Debt}{Number\ of\ shares\ outstanding}$$

Inserting into the formula for expected return

$$E(R_i) = \frac{\left(\frac{Terminal\ Value_1 - Debt_1}{Number\ of\ shares\ outstanding_1}\right) - \left(\frac{Terminal\ Value_0 - Debt_0}{Number\ of\ shares\ outstanding_0}\right)}{\left(\frac{Terminal\ Value_0 - Debt_0}{Number\ of\ shares\ outstanding_0}\right)}$$

Further expanding the formulation

$$E(R_i) = \frac{\left(\frac{\left(\frac{NOPLAT_2\left(1-{Growth_1}/{ROIC_1}\right)}{WACC_1 - Growth_1}\right) - Debt_1}{Number\ of\ shares\ outstanding_1}\right) - \left(\frac{\left(\frac{NOPLAT_1\left(1-{Growth_0}/{ROIC_0}\right)}{WACC_0 - Growth_0}\right) - Debt_0}{Number\ of\ shares\ outstanding_0}\right)}{\left(\frac{\left(\frac{NOPLAT_1\left(1-{Growth_0}/{ROIC_0}\right)}{WACC_0 - Growth_0}\right) - Debt_0}{Number\ of\ shares\ outstanding_0}\right)}$$

Continue the same expansion reveals the formula in its most primal form

$$E(R_i) = \frac{\left(\frac{\left[\left(\frac{\left([Ebit_2 * (1-T_c)] * \left[1 - \frac{\left(\frac{Revenue_2 - Revenue_1}{Revenue_1}\right)}{\left(\frac{Ebit_1 * (1-T_c)}{Invested\ Capital}\right)}\right]\right)}{\left[\left(\frac{E}{V}\right) * R_e + \left(\frac{D}{V}\right) * R_d * (1-T_c)\right] - \left[\frac{Revenue_2 - Revenue_1}{Revenue_1}\right]}\right) - Debt_1\right]}{Number\ of\ shares\ outstanding_1}\right) - \left(\frac{\left[\left(\frac{\left([Ebit_1 * (1-T_c)] * \left[1 - \frac{\left(\frac{Revenue_1 - Revenue_0}{Revenue_0}\right)}{\left(\frac{Ebit_0 * (1-T_c)}{Invested\ Capital}\right)}\right]\right)}{\left[\left(\frac{E}{V}\right) * R_e + \left(\frac{D}{V}\right) * R_d * (1-T_c)\right] - \left[\frac{Revenue_1 - Revenue_0}{Revenue_0}\right]}\right) - Debt_0\right]}{Number\ of\ shares\ outstanding_0}\right)}{\left(\frac{\left[\left(\frac{\left([Ebit_1 * (1-T_c)] * \left[1 - \frac{\left(\frac{Revenue_1 - Revenue_0}{Revenue_0}\right)}{\left(\frac{Ebit_0 * (1-T_c)}{Invested\ Capital}\right)}\right]\right)}{\left[\left(\frac{E}{V}\right) * R_e + \left(\frac{D}{V}\right) * R_d * (1-T_c)\right] - \left[\frac{Revenue_1 - Revenue_0}{Revenue_0}\right]}\right) - Debt_0\right]}{Number\ of\ shares\ outstanding_0}\right)}$$

We see that these variables are contained within the formula for expected return and should therefore, mathematically, influence the expected return. It is therefore a natural starting point to study the effect of these individual variables on $E(R_i)$ and subsequently value. Below we provide an exhaustive list of variables derived from the Gordon growth model (''GGM'').

- EBIT
- Tax rate
- Equity-to-value-ratio
- Debt-to-value ratio
- Interest rate on debt
- Cost of equity
- Invested Capital

- Revenue
- Return on invested capital
- Sales growth
- Weighted average cost of capital
- Net operating profit less adjusted taxes
- Number of shares outstanding

Hence, there is a great number of observable variables originating from corporate finance's fundamental analysis that can, theoretically, have an impact on expected asset returns. This intuition carries over to our research question; *Can the fundamental analysis of the underlying value of a company contribute to well established asset pricing models?*

Previous research has already established that the following four anomalies contribute to asset returns (Fama & French, 2015).

- Firm size (SML-Variable)
- Market price relative to book price (HML-Variable)
- Profitability (RMW-Variable)
- Invested Capital (CMA-Variable)

However, do these variables cover the effect of each variable implied by the GGM? It is clear from the very start that this will be the case for some variable, e.g. invested capital. Invested capital is covered by the CMA-variable, but this variable only measures the amount of capital invested by the firm. It does not, however, measure the return on this capital and will thus exclude the effect of positive investment returns on the asset return. In other words, we find these variables to be quite general, excluding the potential effects of ''smaller'' variables identified through the GGM. Therefore, we aim to test the identified variables after controlling for factors already contained within the Five-Factor model to potentially uncover an increase in the explanatory power of the model. In the case where one of our factors is highly correlated with already contained within the Five-Factor model, we can for example compare the two factors in order to see which one provides more explanatory power than the other, and in such a way potentially introduce alternative factors to those already established.

2.0 Methodology

In general, there are two widely known methodologies applied when trying to identify anomalies (Fama & French, 2008).

i. Sorts of returns on anomaly variables
ii. Regressions

To keep our research methodology in line with Fama and French (1992, 1993, 1996, 2008 and 2015) and thus create the most comparable result as possible, we will apply the method on which these studies are based. This method is described in the early works of Fama and MacBeth (1973), where they use anomaly variables to explain the cross-section of average returns (Fama & French, 2008).

2.1 Cross-section regressions and the Fama and MacBeth (1973) methodology

In true Fama and French fashion, we set to study all stocks registered on NYSE at the same point in time. Hence, we are dealing with a cross-sectional data set and need to apply techniques accordingly (Woolridge, 2011).

2.1.1 Cross-section regressions

There are two steps required when estimating regressions on a cross-sectional data set (Woolridge, 2011). The first step requires you to run a regular time-series regression for each point in time, *t*, for each asset, *i*. For illustrative simplicity, we represent a regression equation with one factor, *F*.

$$R_{i,t} = a_i + \beta_i * F_t + \varepsilon_{i,t}$$

We run this equation and obtain the estimates for all β. This will be the measure of sensitivity of return, $R_{i,t}$, on the factor, F_t, and can be calculated using ordinary least squares (‘’OLS’’).

Step two is to run the actual cross-sectional regression, which will differ slightly from the original time-series regression in that we use the beta-estimates obtained from step one as the explanatory variables on the right-hand side of the equation, rather than the factor itself. Now using expected return on asset, we get the following equation.

$$E(R_{i,t}) = \gamma_i * \beta_i + \alpha_i$$

Hence, the goal is to estimate the γ_i which can be interpreted as the ''risk premium'' for exposure to that factor. This can be either positive or negative, significant or insignificant. If it proves to be statistically significant, then we have found support that the factor *might* contribute to higher, although we must be very careful with such conclusions.

2.1.2 Fama and MacBeth (1973) approach

One major issue with cross-sectional analysis of financial data is that the error terms tend to be correlated across assets. For example, if one firm oil-prices rise then the stock price of oil-producing firms will also rise. Since we do not include ''oil-prices'' as a factor in our model, this will be picked up by the error term, making the error-terms across oil-producing firms correlated.

The Fama-MacBeth (1973) provides a solution to this problem that differs very slightly from the 'regular' approach. The first step remains the same, i.e. we run a time-series regression for each point in time, *t*, for each asset, *i*, and obtain the beta-estimates. Assuming that we have a total of n assets and m factors, we have to run the following time-series regression in order to capture the exposures to each factor, *j*.

$$R_{1,t} = \propto_1 + \beta_{1,F_1} F_{1,t} + \beta_{1,F_2} F_{2,t} + \cdots + \beta_{1,F_m} F_{m,t} + \varepsilon_{1,t}$$
$$R_{2,t} = \propto_2 + \beta_{2,F_1} F_{1,t} + \beta_{2,F_2} F_{2,t} + \cdots + \beta_{2,F_m} F_{m,t} + \varepsilon_{2,t}$$
$$\ldots\ldots$$
$$R_{n,t} = \propto_n + \beta_{n,F_1} F_{1,t} + \beta_{n,F_2} F_{2,t} + \cdots + \beta_{n,F_m} F_{m,t} + \varepsilon_{2,t}$$

As before, we store the *m* amount of beta-estimate and regress these as independent variables against the return on asset. However, now we compute cross-sectional regressions at each point in time, *t,* to T.

$$R_{i,1} = \gamma_{1,0} + \gamma_{1,1}\hat{\beta}_{i,F_1} + \gamma_{1,2}\hat{\beta}_{i,F_2} + \cdots + \gamma_{1,m}\hat{\beta}_{i,F_m} + \varepsilon_{i,1}$$

$$R_{i,2} = \gamma_{2,0} + \gamma_{2,1}\hat{\beta}_{i,F_1} + \gamma_{2,2}\hat{\beta}_{i,F_2} + \cdots + \gamma_{2,m}\hat{\beta}_{i,F_m} + \varepsilon_{i,2}$$

$$\ldots.$$

$$R_{i,T} = \gamma_{T,0} + \gamma_{n,1}\hat{\beta}_{i,F_1} + \gamma_{n,2}\hat{\beta}_{i,F_2} + \cdots + \gamma_{n,m}\hat{\beta}_{i,F_m} + \varepsilon_{i,T}$$

At this point, we will have *m* estimates of γ (+1 if we choose to include the constant term). However, in order to calculate the risk premium for each of the risk factors, we need some kind of average of $\hat{\gamma}$ for each risk factor, F. Assuming that the error terms, $\varepsilon_{i,t}$, are independent and identically distributed (''I.I.D''), we can calculate the risk premium for the *m*-th factor, F using the following equation;

$$Risk\ Premium\ for\ factor\ m = \frac{\sum_{i=1}^{n} \gamma_{i,m}}{T}$$

2.2 Testing for significance

The main goal of this thesis is to identify variables that potentially can influence asset returns. Essentially, this translates to statistically significant risk premiums, $\hat{\gamma}$. In other words, we want to test if the different risk premiums calculated are statistically different from 0, thus implying that the risk factor potentially influence asset returns.

Before we can start testing for significance, we need to obtain the t-statistics. Since we are working with γ-estimates that are averaged over time, we must compute the Fama and MacBeth (1973) variant of the t-statistic (See Fama (1965); Blume (1970); and Officer (1971) for justification).

$$t(\bar{\hat{\gamma}}_m) = \frac{\bar{\hat{\gamma}}}{\frac{s(\hat{\gamma}_m)}{\sqrt{n}}}$$

Since we are to test the risk premiums, γ, for statistical significance, the hypothesis of main interest will be as follows.

$$H_0: \bar{\hat{\gamma}}_j = 0$$

$$H_a: \bar{\hat{\gamma}}_j \neq 0$$

Hence, we will apply the approach suggested by Fama and MacBeth (1973) to check our factors for statistical significance. It is, however, important to note that further testing may occur as we progress our reading and understanding of the topic.

2.3 The Variables and potential anomalies

In the introduction to our thesis, we identified potential variables that may or may not prove to be anomalies. Our intention is to test all these variables after controlling for the five factors already contained within the Five-Factor Model. However, if it occurs that one or more of the variables prove to be unobservable or incomplete, these might be removed from the final version of our thesis. We will also run several variants of the model, changing up the variables included, to see which one has the better performance. An example of such a ‘’tweak’’ is the SML-variable included in the original Three-Factor model, which we might split up into two; e.g. make one equity/value-variable and one debt/value-variable.

Further on, to make the variables more comparable across a wide range of different companies and sectors, we will standardize each set of variables by converting them into ratios/multiples. This also seems to be the standardized approach used by researchers studying the effect of anomalies (see Bondt & Thaler (1985), Basu (1983), Jegadeesh & Titman (1993) and Basu (1983)).

2.3.1 An exhaustive list of variables

As previously mentioned, we will test several different models where we include or exclude various variables to see which one has the better performance. Therefore, we cannot at this point explicitly state which variables we will focus the most heavily on as we do not know which are significant or not, or which ones that are to highly correlated with already existing variables in the Five-Factor model. For this reason, we will simply provide an exhaustive list of variables which may, or may not be included in the final model.

- $\frac{EBIT}{Enterprise\ value}$
- Tax rate
- $\frac{Equity}{Enterprise\ value}$
- $Leverage\ multiple = \frac{Net\ interest\ bearing\ debt}{Enterprise\ value}$

- Interest rate on debt
- $\frac{Invested\ Capital}{Enterprise\ value}$
- $\frac{Revenue}{Enterprise\ value}$
- Return on invested capital in %
- Sales growth in %
- $\frac{NOPLAT}{Enterprise\ value}$
- Number of shares outstanding

2.3.2 Circularity issue using WACC as an independent variable

The weighted average cost of capital can prove to create a circularity issue within our regression. This can be observed in the definitional equation for WACC:

$$WACC = \left(\frac{E}{V}\right) * R_e + \left(\frac{D}{V}\right) * R_d * (1 - T_c)$$

R_e represents the cost of equity for the firm, which in turn is by practitioners within corporate finance calculated using the classical SLB-model. Hence, we cannot use WACC as a variable because the formula for this variable contains the independent variable we intend to estimate. I.e. to avoid the circularity issue, we intend to ignore WACC.

3.0 Data

As our research paper is going to be built upon the work of Fama and French, we will be using some of the same sources for data collection. Further on, Kenneth R. French has created a substantial data base which can potentially be a vital source of data for our thesis ("Kenneth R. French - Data Library", 2018).

The left-hand side (''LHS'') of our regression is going to be monthly percentage returns, including dividend and capital gains (with appropriate adjustments for capital changes, such as splits and stocks). The data needed for this part can be downloaded from the Center of Research in Security Prices of the university of Chicago (''CRSP''). CRSP provides a list giving a detailed description of distributions made to shareholders of a security, such as; cash dividends, stock splits, spin-offs, rights offerings, liquidation payments, etc. ("Stock Data Structure | CRSP - The Center for Research in Security Prices", 2018). The CRSP also has data on the number of shares outstanding. This can be divided into two parts. One is a primary share observation (taken directly out of the company annual or quarterly reports), the second is supplementary observations derived from all the distributions affecting shares outstanding. There are also programs available to map observations to time series used to calculate market capitalization. CRSP have a merged database with Compustat which is going to be helpful for analyzing and extracting relevant data ("CRSP/Compustat Merged Database | CRSP - The Center for Research in Security Prices", 2018).

For all the other variables listed under "Our thesis" we will be using available terminals such as Bloomberg and Eikon.

Thomson Reuters (Eikon) database contains information about over 6000 global companies, and all the data is quality controlled and collected such that the information is standardized and comparable ("ESG Research Data", 2018). The database gives the opportunity to compare corporate data against peers and give insights into corporate structures and performance of the biggest companies for investment purposes. The data needed for the research can then be customized at the platform ready for download via the "Datastream sample sheets".

Bloomberg also offers tools used by professional analysts to analyze industries and competitors in addition to information about the capital structure of firms ("Bloomberg Terminal | Bloomberg Professional Services", 2018). The technology can be utilized to assess the fair value of corporations based on historical performance ("Investor Relations | Bloomberg Professional Services", 2018).

4.0 Reference review

The literature and research in this field of study can be dated back to the early 1960`s, and it is striking how little knowledge there was about risk and return before the introduction of the CAPM. Theories about risk preferences were not developed until the 1940`s and 1950`s. Starting with the CAPM we will present a review of the literature in relevance to our area of research.

The CAPM was for long recognized as the most reliable asset pricing model, that is until Fama and MacBeth, (1973) used the well-known "two-parameter" portfolio model to test the relationship between the average return and risk for common stocks traded on the New York Stock Exchange. The results implied that there is a linear relationship between the risk and return, and no evidence supporting that no other measure of risk or factor in addition to portfolio risk, systematically affects average returns.

A few decades later Fama and French (1992) introduced two easily measured variables, size and book-to-market equity. This was an extension of the "two-parameter" portfolio model mentioned above. The aim was to capture the cross-sectional variation in average stock returns associated with market beta, size leverage book-to-market equity and earnings-price ratios. To increase the explanatory power of the model capturing the size and book-to-market equity, Fama and French (1993) presented another model identifying five risk factors in the returns of stocks and bonds. Three of them were stock market factors which formed the now famous Fama French Three-factor model. This is recognized as a great breakthrough in the field of finance.

Fama and French (1996) continued testing the Three-factor model and observed that the anomalies almost disappeared except the continuation of short term returns. The results were consistent with the rational ICAPM and APT asset pricing, so possible explanations could have been irrational pricing and data problems. The authors admit that even though the Three-factor model is a good model, it does not capture all the factors e.g. the continuation of short term returns.

Fama and French (2008) continue to study anomalies, this time focusing on the returns associated with net stock issues, accruals and momentum. To explore these return anomalies cross-section regressions were used on microcaps, small stocks and big stocks. An interesting result is that the regressions seems to have unique information about future returns. One important note is that variables that predict expected cash flows also predict returns, but they do not tell us how much of the variation in expected returns is due to risk and how much is caused by the mispricing. These results imply that it worth studying these anomalies, given that they are still producing fresh insights.

The latest attempt to capture the anomalies that are not explained by the CAPM is were Fama and French (2015) extends the Three-factor model introducing a five-factor asset pricing model. The three-factor model was criticized for being an incomplete model because it did not capture the variation in average returns associated with profitability and investment, hence the two factors added to the new model is profitability and investment. Although there is a lot to extract from this study, there is no clear answer on if the five-factor model is a more superior model.

As we mentioned in the earlier sections there is a lot of assumptions underlying which all this research is built upon. Black (1972) have explored the nature of capital market equilibrium, experimenting with some of the underlying assumptions deriving the capital asset pricing model. One such assumption that have been tested is that all investors have the opportunity to invest and borrow at a risk-free rate. Then it has been assumed that investors are allowed to only invest i.e. take long positions in the risk-free asset neglecting the opportunity of short positions (borrowing). The author concludes that the empirical results are consistent with a market equilibrium in both cases. This approach may not be suitable for other assumptions in different situations, but the article can be used to help us develop alternative strategies to counter other relatable problems.

Even though it already exists a lot of research in this field, there is still a lot left to uncover. Fama and French's research is still up to date and taught in every business school today. Fama and French (1996) states," There is an important hole in our work", and no matter how good a statistical model is there is always a positive probability that it can be improved further.

5.0 Reference list and data sources

5.1 Reference list

Banz, R. (1981). The relationship between return and market value of common stocks. *Journal Of Financial Economics*, *9*(1), 3-18. http://dx.doi.org/10.1016/0304-405x(81)90018-0

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks. *Journal Of Financial Economics*, *12*(1), 129-156. http://dx.doi.org/10.1016/0304-405x(83)90031-4

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal Of Business*, *45*(3), 444. http://dx.doi.org/10.1086/295472

Blume, M. (1970). Portfolio Theory: A Step Toward Its Practical Application. The Journal Of Business, 43(2), 152. http://dx.doi.org/10.1086/295262

Bondt, W., & Thaler, R. (1985). Does the Stock Market Overreact?. *The Journal Of Finance*, *40*(3), 793. http://dx.doi.org/10.2307/2327804

Clarke, K. (2005). The Phantom Menace: Omitted Variable Bias in Econometric Research. *Conflict Management And Peace Science*, *22*(4), 341-352. http://dx.doi.org/10.1080/07388940500339183

Fama, E. (1965). The Behavior of Stock-Market Prices. The Journal Of Business, 38(1), 34. http://dx.doi.org/10.1086/294743

Fama, E., & French, K. (1992). The Cross-Section of Expected Stock Returns. *The Journal Of Finance*, *47*(2), 427-465. Retrieved from http://www.jstor.org/stable/2329112

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal Of Financial Economics*, *33*(1), 3-56. http://dx.doi.org/10.1016/0304-405x(93)90023-5

Fama, E., & French, K. (1996). Multifactor Explanations of Asset Pricing Anomalies. *The Journal Of Finance*, *51*(1), 55. http://dx.doi.org/10.2307/2329302

Fama, E., & French, K. (1996). The CAPM is Wanted, Dead or Alive. *The Journal Of Finance*, *51*(5), 1947. http://dx.doi.org/10.2307/2329545

Fama, E., & French, K. (2008). Dissecting Anomalies. *The Journal Of Finance*, *63*(4), 1653-1678. http://dx.doi.org/10.1111/j.1540-6261.2008.01371.x

Fama, E., & French, K. (2015). A five-factor asset pricing model. Journal Of Financial Economics, 116(1), 1-22. http://dx.doi.org/10.1016/j.jfineco.2014.10.010

Fama, E., & MacBeth, J. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal Of Political Economy*, *81*(3), 607-636. http://dx.doi.org/10.1086/260061

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal Of Finance*, *48*(1), 65. http://dx.doi.org/10.2307/2328882

Koller, T., & Goedhart, M. (2015). *Valuation*. Hoboken: Wiley.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review Of Economics And Statistics*, *47*(1), 13. http://dx.doi.org/10.2307/1924119

Markowits, H. (1952). Portfolio Selection. *The Journal Of Finance*, *7*(1), 77-91. Retrieved from http://www.jstor.org/stable/2975974

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. Journal Of Financial Economics, 108(1), 1-28. http://dx.doi.org/10.1016/j.jfineco.2013.01.003

Officer, R. (1971). The Variability of the Market Factor of the New York Stock Exchange. The Journal Of Business, 46(3), 434. http://dx.doi.org/10.1086/295551

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal Of Finance*, *19*(3), 425. http://dx.doi.org/10.2307/2977928

Titman, S., Wei, K., & Xie, F. (2004). Capital Investments and Stock Returns. Journal Of Financial And Quantitative Analysis, 39(04), 677. http://dx.doi.org/10.1017/s0022109000003173

Wooldridge, J. (2011). *Econometric analysis of cross section and panel data*. Cambridge, Mass.: MIT.

Wooldridge, J. (2016). *Introductory econometrics*. Boston, Mass. [u.a.]: Cengage Learning.

Lo M. (2013) Market Efficiency Hypothesis. In: Lee CF., Lee A. (eds) Encyclopedia of Finance. Springer, Boston, MA

5.2 Data sources

Bloomberg Terminal | Bloomberg Professional Services. (2018). *Bloomberg Professional Services*. Retrieved 13 January 2018, from https://www.bloomberg.com/professional/solution/bloomberg-terminal/

CRSP/Compustat Merged Database | CRSP - The Center for Research in Security Prices. (2018). *Crsp.com*. Retrieved 13 January 2018, from http://www.crsp.com/products/research-products/crspcompustat-merged-database

ESG Research Data. (2018). *Thomson Reuters*. Retrieved 13 January 2018, from https://financial.thomsonreuters.com/en/products/data-analytics/company-data/esg-research-data.html

Investor Relations | Bloomberg Professional Services. (2018). *Bloomberg Professional Services*. Retrieved 13 January 2018, from https://www.bloomberg.com/professional/expertise/investor-relations/

Kenneth R. French - Data Library. (2018). *Mba.tuck.dartmouth.edu*. Retrieved 13 January 2018, from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Stock Data Structure | CRSP - The Center for Research in Security Prices. (2018). *Crsp.com*. Retrieved 13 January 2018, from http://www.crsp.com/products/documentation/stock-data-structure