SCHOOL OF
ECONOMICS AND
MANAGEMENT

Does Quality Matter?

An Empirical Study of the Size Premium and its Relation to Firm Quality

by

Lovisa Jalrup

Sawan Patel

May 2020

A thesis presented for the degree of

Master's in Finance

Supervisors: Hossein Asgharian and Duc-Hong Hoang

Department of Economics

Lund University

Abstract

The purpose of this study is to investigate if there is any size effect in the Swedish stock market between April 2010 and December 2019, and if controlling for firms' quality improves the performance of a size-based investment strategy. The risk premium of firms with smaller market value of equity has since its discovery been under heavy scrutiny. Recent studies suggest that by quality-control, the size-based investment strategy can be improved. It is therefore highly relevant for the investor to understand how quality affects such a strategy. By employing a double sorting portfolio construction by size and quality as well as a cross-sectional regression using the Fama-MacBeth two-step regression, we examine if a small-minus-big portfolio yields a positive excess return and the interaction of quality and the size risk factor. We find that there is no size effect without any quality-control, in the Swedish stock market. However, in conjunction with the quality metrics, size is a relevant risk factor in asset pricing. In conclusion, by constructing the portfolio of high-quality firms, in terms of the Return on Assets or the credit rating of a Merton-based Credit Risk Model, the size-based investment strategy can be improved.

Keywords: Investment strategy, Small-Minus-Big (SMB), size effect, firm quality, Fama-MacBeth two-step regressions.

Acknowledgements

We would like to thank Hossein Asgharian and Duc-Hong Hoang for their quick, helpful, and constructive feedback throughout the paper.

Lovisa Jalrup & Sawan Patel

Contents

Definitions

TSR: Time Series Regression

CSR: Cross Sectional Regression

OLS: Ordinary Least Squares

GLS: General Least Squares

MV: Market Value of equity

BM: Book-to-market value

ROA: Return on Assets

SMB: Small minus Big

QMJ: Quality minus Junk

FF3: Fama-French Three Factor Model

EW: Equally-Weighted

VW: Value-Weighted

WTP: Willingness to pay

1. Introduction

Banz (1981) discovers an anomaly of stock market equity returns known as the size effect. The anomaly is the tendency that smaller firms, in terms of the market value of equity, outperform larger firms, yielding a premium for smaller firms. However, the existence of the size effect has come under heavy scrutiny. Results from later studies indicate that the size effect from time to time disappears and reappears, and even sometimes is negative under different circumstances. On the other hand, other papers indicate that investors can increase the robustness of the size premium on their portfolio by controlling for firm quality. In theory, the size premium exists since smaller firms can be considered riskier than bigger firms, making the size premium a form of risk compensation. Empirical findings show that higher quality of firms makes for safer investments, but also yields higher returns. Hence, an investor, by adding quality metric to the investment decision, can sort out low-quality firms that potentially distorts the size premium, from their portfolio. The aim of this paper is to investigate if there is any size effect in the Swedish stock market, and if controlling for firms' quality improves the performance of a size-based investment strategy.

It is highly relevant for investors to understand the size effect and its impact on the return on investment. More generally, such a factor can be systematically used to form an investment strategy. There is a substantial interest in mutual funds that invest in small cap firms in the Swedish market. There are around 80 mutual funds sold to Swedish investors that mainly focus on small cap firms (Morningstar, 2020). An investment strategy based on the size effect can be a portfolio in which the investor goes long in small stocks and short in big stocks (SMB). The main reason for the importance of the quality-control in a SMB portfolio is a risk of going long in small firms of poor quality and shorting large high quality firms, the two anti-poles distorts the size effect. This is based on the fact that high quality firms (good firms) tends to outperform low quality firms (bad firms) (Asness, Frazzini, and Pedersen, 2019).

According to the capital asset pricing model (CAPM), the cross-section of expected stock returns can be explained by a linearly positive relationship to the firms' systematic market risk, measured by the β (market beta) (Fama and MacBeth, 1973). Banz (1981) introduces the size effect but does not draw any clear conclusion of whether the size effect is a factor on its own or rather is a proxy for other factors correlated to the size of companies. Either way, Banz (1981) questions if there could potentially be something missing from the traditional CAPM specification. Fama and French (1992) confirms, that the beta in CAPM is not enough to explain the cross-section of expected stock returns. Both Banz (1981) and Fama and French (1992) argue that their discovery is not necessarily an evidence against the market efficient hypothesis, but rather that the traditional asset pricing model needs to be more correctly specified. By adding both a size factor and a book-to-market factor to the asset pricing model specification, in addition to the market factor, Fama and French (1992) finds more consistency in the results from the re-specified asset pricing model.

However, size factor seems to be very weakly present over a longer time period. Asness et al. (2018) finds the disputed size effect to be consistent across time, but not unless the size has been controlled for quality. Banz (1981) was early to indicate that the size effect was perhaps not a very stable factor on its own, so by adding a controlling factor such as quality, Asness et al. (2018) find the size factor strongly resurrected throughout all the time periods tested (July 1926 to December 2012), including the periods where the size factor historically was insignificant. The findings are strong indications that the quality of a firm is of importance for the existence of size premiums. There are several theories as to why quality is related to the size factor. Asness et al. (2018) finds that size interacts strongly with quality variables, and thereby quality is a factor which is helpful in the description of the size and expected returns relationship. Further, quality may be a very broad company specific characteristic, but Asness et al. (2018) define quality as any characteristic of a firm that an investor has a higher willingness to pay for, all else equal. Once there is a control for the quality, they find that small stocks outperform the big stocks when companies are comparable in the level of quality. Thereby, the quality-control returns the size effect to be a positive factor to the expected returns in the cross-section of asset returns. Once a control for quality is in place the investor could simply go long in only small firms of high quality and short the large low quality firms. They conclude that, no matter which quality control is used, the size effect emerges.

In this study, we are looking into the Swedish market, and use all the actively listed stocks on the Stockholm Stock Exchange over the time period 2010-04-01 to 2020-01-01. The time period is limited based on the availability of data. We use four different quality variables; ROA, Leverage, SmartRatio Credit Risk model and Structural Credit Risk model. Hagel III, Seely Brown, and Davison (2010) argue that return on assets (ROA) is a good measure of company performance and George and Hwang (2010) and Penman, Richardson, and Tuna (2007) finds a negative relation between leverage and return. High leverage indicate higher risk for bankruptcy, therefore low leverage is generally considered as safer for an equity investor and fits in to the definition of quality. The SmartRatio Credit Risk model and the Structural Credit Risk model are both measures of credit ratings provided by Thomson Reuters, however, they are complements to each other as they are constructed based on different information sets and models, and therefore treated as two separate quality metrics. The Structural Model is based on the Merton Model to estimate probability of default and reflects quality in sense of business- (volatility) and financial risk (leverage). The SmartRatio Model is based on the firms' financial ratios, better ratios indicates better firm quality in terms of performance. The data for the two credit risk models are updated daily, therefore these makes for dynamic proxies of quality. The time period of the sample (2010-2019) is limited based on the availability of data for the SmartRatio- and Structural Model. The quality variables that is used in our study is further motivated by the easy availability of the information to investors and its intuitive economic interpretation.

We apply two different methods in search for the size premium. The first method we apply in order to evaluate the size effect is by creating a SMB portfolio and compute the monthly average return. We also perform double sorting on size and quality, in order to see how quality interacts with size in a similar manner as Asness et al. (2018) by computing the monthly average return. The second method we apply is in order to see if risk factors, such as size, can explain the cross-sectional excess returns. i.e. if there is a risk premium for size across firms and if it is affected by quality after controlling for market beta. We use a two-step approach, originally presented by Fama and MacBeth (1973), where in the first step we run time-series regressions (TSRs) with a rolling window of three years of monthly data to estimate the individual market betas. This is followed by the second step of running cross-sectional regressions (CSRs) on the betas along with additional explanatory variables, such as size, book-to-market and quality, for the cross-section of excess returns.

In both applied methods, the data sample contains companies that are widely varied in terms of size and quality. The main contribution relates to the control for quality. We are able to include many small stocks with credit ratings (which is typically less common) by the use of the SmartRatio and Structural Model. If we would have used the credit ratings from one of the tree largest credit rating agencies (Moody's, Fitch, or S&P), the sample would include predominantly large companies. Furthermore, our focus is on the risk premium for firm size.

We do this to establish if firm size is a relevant risk factor to systematically be used in the investment decision, unlike most research that focus on the abnormal returns of the size-based investment strategy.

The main result of this study is that the size effect does not plainly seem to exist in our sample, but by including a quality variable, the size effect appears. Thereby, quality is of importance for the investor following the size-based investment strategy. Which of the quality measurements the investor chooses to look at does matter for the significance of the size premium and the SMB average return. By including ROA and the Structural credit risk model in the Fama and MacBeth (1973) two-step regressions, a negative risk premium appears for firms with larger market value of equity, and positive return for the quality variable. The monthly average excess return of SMB portfolios shows that using good quality firms yields a positive size premium while using low-quality firms yields negative returns. In other words, an investor can improve the size-based investment strategy by forming it based on high quality firms.

Following this introduction is a summary of the earlier research in Section 2, after which the theoretical background is presented in Section 3. The data and methods used in this paper is found in Section 4, followed by the presentation of results and analysis in Section 5. Finally, the conclusion and some suggestions on further research concludes this paper in Section 6.

2. Earlier Research

This chapter reviews the earlier research on the size effect as well as on the quality effect. It includes the theoretical background on the size effect and some critique on its existence. Furthermore, earlier research on qualities effect on returns and its effect on the size effect is reviewed. Lastly, the contribution of our study to the previous research is discussed.

2.1 Size effect

One prominent critic against the size effect, presented by Banz (1981), in particular with regards to the framework of Fama and French, is Black (1993). He points out that Fama and French (1992) show that size could explain cross-sectional returns during overlapping sample time with Banz (1981). However, for the period after Banz's discovery, they do not find the size effect at all. This is aligned with several studies showing that the size effect has since its first discovery somewhat vanished. Black (1993) argue that the size effect is just a product of data mining, meaning expected returns is solely based on historical patterns discovered by chance, and the lack of theory can explain the later absence of the size effect. On the other hand, Berk (1995) argue that size is a highly relevant risk factor and should not be considered as an anomaly, but rather be included in the asset pricing models. The reasoning is as follows; suppose there are two firms with the same operating size, meaning same expected end-of-period cash flow, but have different risk from each other, the riskier firm will have lower beginning-of-period equity market value because of discounting and hence higher expected return. Berk (1995), however, refrains to conclude what the size is proxy for, he finds no evidence of non-risk-adjusted size measures, such as book value of assets, to have a relation to return.

Schwert (2002) describes anomalies as those empirical results that contradicts conventional asset pricing models and could occur because of market inefficiency or misspecified asset pricing model. His study, however, also finds that size effect (and other anomalies) has weakened or disappeared after its discovery when looking at US stocks between 1962 and 2001. He argues that after the discoveries, investors begin to implement investment strategies that utilizes the arbitrages, essentially trading it away. Crain (2011) reviews previous research and finds varied results. When size is found, it can be explained by some background factor, such as higher liquidity risk and less information on small firms leading to higher returns. Other results show that size effect is concentrated to the very smallest firms and is nonlinear. On the other hand, some of the results shows no size effect at all. He concludes that since the relevant risk that size is a proxy for is not consistent, the real risk factor might not anymore be correlated with size. Chordia, Subrahmanyam, and Tong (2014) find that reduced tick size leads to a smaller size premium. They believe this could mean that size effect arises from liquidity risk but disappears as modern technology allows for lower transaction costs. Horowitz, Loughran, and Savin (2000) looks at the size effect between 1980 to 1996 and find no significant relation between firm size and asset returns. Israel and Moskowitz (2013) find no significant abnormal returns of SMB over wide periods between 1926 and 2011. Gompers and Metrick (2001) finds that the size effect has disappeared in the US at least between 1980 and 1996 due to institutional investors continued to demand big firms which increased the returns on those assets.

Reinganum (1983) finds the size effect but concentrated to January, which has given the size effect the name January effect in the literature. He argues that the effect arises from that investor sell of their assets for tax-loss reasons in December and repurchase in January, in particular poor performer of the previous year. Gu (2003) however finds that the January effect to be inconsistent. It has followed a downward trend in some periods and index, and positive trend in others. The size effect is most apparent in indexes consisting of large firms, indicating that the January effect is in fact not a size effect.

Asness et al. (2018) also examine if there is size effect, before controlling for quality, under different settings and finds that it is most significant in January in the US for the years between 1926 and 2012, but insignificant for all other months. It is significant for the period 1957 to 1979 but disappears after that. They also look at the global, the Pacific and the European market and find no significant size effect. They use a double sorting method to divide the firms into portfolios based on size and quality. They find that by forming SMB portfolio from the intersections of size and quality, no size effect appears amongst the worst quality firms, however, positive and significant for all other quality quintiles.

Dichev (1998) acknowledges that the absence of size effect and examine if size could be a proxy for bankruptcy risk, however, no such relation is found. Instead he shows that firms with higher bankruptcy risk yields less average return since 1980.

2.2 Introduction of Quality

Campbell, Hilscher, and Szilagyi (2008) research if there is a return premium for investors holding financially distressed stocks. The main relationship explored is that between the pricing of stocks and the credit risk of the stocks. High credit risk is defined commonly as a high probability of default within the next 12 months. The total period covered in their study is from January 1963 to December of 2003, and the geography covered is the US market. The main results by Campbell, Hilscher, and Szilagyi (2008) is that the stocks presenting a high credit risk, yields anomalously low average returns. In addition, these also tend to have higher standard deviations and higher market betas in comparison to the companies that have a low credit risk. Further, distressed firms seem to be more sensitive to an increase in market volatility or in the risk aversion of investors, indication that the poor-quality stocks are abandoned for the higher quality stocks in times of general market distress (Campbell, Hilscher, and Szilagyi, 2008). The hypothesis that the size or value effect, from the three-factor model developed by Fama and French (1992), is capturing the compensation of financial distress, is inconsistent with the results according to Campbell, Hilscher, and Szilagyi (2008). The inconsistency is attributed to the distress effect being the same across different sizes, even if the small firms tend to be in more severe financial distress than the larger firms. There are high loadings on a SMB factor indicating that the credit risk control variable results in even worse returns for the low quality (high credit risk) stocks.

George and Hwang (2010) adds leverage, in addition to financial distress, as a quality variable explaining the cross-section of returns. The returns are negatively correlated to the leverage of a company. Similarly, Penman, Richardson, and Tuna (2007) finds a negative relationship between leverage and returns. The data, from both George and Hwang (2010) and Penman, Richardson, and Tuna (2007), consists of US companies from the early 1960s to the early 2000s. Adding financial risk, such as an increase in leverage, should rationally give higher returns, since higher risk should reward the investor with a higher expected return. According to Penman, Richardson, and Tuna (2007), this basic understanding of the risk-return relationship is violated when their results show a negative relationship between high leverage firms and returns. The authors attribute this to three possible explanations, either leverage is very wrongly priced, that these are just sample-specific results, or that leverage is capturing another risk factor in a distorted way. The last bringing about a warning of general misspecification of asset pricing models. George and Hwang (2010) on the other hand, argues against mispricing based on their results of leverage affecting the asset returns similarly for companies, no matter the intensity of analysts' reports of the companies. The main argument to the negative relationship between leverage and returns are explained by these authors as underlying financial distress costs differences between the firms. The explanation being that low leverage means higher distress costs as a result of a larger exposure to the systematic risk, unlike the highly leveraged companies. The probability of default is also low in the companies of low leverage, so the low leverage implies a higher financial distress costs, and thereby negatively related to expected returns. The investors in the equity market knows that the capital structure that minimizes the financial distress costs, hence the negative relation between leverage and returns is not inexplicable (George and Hwang, 2010).

Most recently, Kyosev et al. (2020) disputes the notion of distress risk being the cause between quality variables and their relationship to return premiums. The authors attribute the quality premium to mispricing rather than the cause being the distress risk. These results are based on data from a larger geographical region; US, Europe, Japan, and emerging markets, over the time period of mid 1980s to year end 2015. The paper focus on accounting ratios and variables, and therefore exclude financial institutions. The accounting variables are those commonly associated with quality measurements, and the predictiveness of these variables to the stock returns. Only the variables that indicate the growth of earnings for a company are shown to be of relevance for the stock return prediction. These findings are also affirming the result of Novy-Marx (2013) which is that the more profitable, as measured by gross-profits normalized by the asset base, also have higher returns.

Asness, Frazzini, and Pedersen (2019) thoroughly review what makes an investor have a higher willingness to pay (WTP) for certain stocks based on quality characteristics. The authors find three main features, namely; profitability, growth of profits, and safety (stocks of a lower return requirement are deemed safer), that investors seem to have a higher WTP for. However, the price of a high-quality stock is not that much higher compared to a lower quality stock. The strategy of going long in high quality stocks and shorting low-quality stocks (quality minus junk, or QMJ for short) does yield high risk-adjusted returns. The results hold true for the 25 countries part of the study over the time period 1957 to 2016.

Asness et al. (2018) examine if the size effect resurfaces by controlling for a wide range of quality metrics. As quality metrics they use the Quality-minus-Junk factor from Asness, Frazzini, and Pedersen (2019) to show that quality can help to explain SMB returns and creates abnormal returns. Furthermore, Asness et al. (2018) include profitability, growth, safety, payout, credit rating, investment factor etc. hedging against the low quality in any kind of quality metric helps to increase and make the size effect strongly significant. It becomes more consistent in all industries, both in the US and worldwide, and over time. Asness et al. (2018) uses double sorting method and examines the average monthly returns. They also look at the intercept as the abnormal return from regressing SMB portfolio on the market return, HML, momentum portfolio and the quality portfolios. With the latter method they also find that size effect is not only concentrated to January.

Our contribution primarily lies in performing the Fama and MacBeth (1973) two-step approach with a quality-control. Additionally, we look for the size effect in a newer data set, whereas most previous research focus on the long-term historical return in the 80's and 90's. If the size effect exists in modern time stock market is highly relevant for the current investors rather than the historical records. Furthermore, as opposed to previous researchers who uses single quality metrics such as various financial ratios and information about growth, payouts, investment rates etc., we also have the two daily updated credit ratings (SmartRatio- and Structural Models are further explained in Section 4.1) that enables us to include a larger number of small firms in our sample. It must also be noted that we rely on that Thomson Reuters, who creates the credit ratings used in this paper, make accurate and correct modeling to be used as proxy for quality.

3. Theoretical Models

This chapter introduces the main theories of cross-sectional expected returns in financial economics, capital asset pricing models and the development into the Fama-French Three Factor model, the latter being based on the size effect. We also introduce the theory on one of the most conventional modelling for credit risk, the Merton Model, which forms the basis for the Structural credit risk model.

3.1 Asset Pricing Models

Expected stock returns are in equilibrium explained according to a single index model (such as CAPM), developed by Sharpe (1964), Lintner (1965) and Mossin (1966) (summarized by Bodie, Kane, and Marcus (2014)). In a CAPM-framework, expected returns on any portfolio of risky assets are explained by the sensitivity to a systematic factor measured by the variable β. The returns are thus a reflection of the portfolio risk, where higher risk should *ex-ante* yield higher expected returns. Risk is measured by the volatility of the asset, in particular it is only the undiversifiable systematic risk that is priced, namely the asset volatility in relation to the systematic factor volatility (see Eq. 3.1), i.e. the beta between the systematic factor and stock return. The relevant systematic factor, which CAPM assumes to solely explain asset return, is the market portfolio that is a fully diversified weighted portfolio of all risky assets. In CAPM, the market portfolio is the general factor for any macroeconomic variable that could or should affect stock prices (Bodie, Kane, and Marcus, 2014). R_i, in Eq. 3.1, is asset i's return in excess of the risk free rate, and R_m is market portfolio return in excess of the risk-free rate.

$$\beta_i = \frac{cov(R_i, R_m)}{var(R_m)} \tag{3.1}$$

If CAPM holds, the excess returns for period t should be equal to what is given by Eq. 3.2, i.e. only be defined by the beta and the market excess return.

$$E(R_{it}) = \beta_i E(R_{mt}) \tag{3.2}$$

An investor is thus not compensated for any idiosyncratic or unsystematic risk that arises when not holding a fully diversified portfolio, since such an investor could just diversify away that risk. When the *ex-post* return for any asset deviates from what is expected using CAPM, the difference is considered as the unexpected return, however, such return is firm specific but should on average equal to zero according to CAPM.

Fama and French (1992) find that beta in the CAPM framework lacks explanatory power for expected stock returns, therefore they investigate other potential risk factors and come to the conclusion which has somewhat changed the industry standard for what an asset pricing model should look like. Most prominent is the addition of the size where there is a negative relation between firm size (MV) and the average stock return. Furthermore, they argue that book-to-market (BM) of equity value has strong positive relation of average stock returns. From this, a conventional asset pricing model is formed, namely the Fama French Three Factor Model (FF3). In the study of Fama and French (1992), the specification for the the cross-sectional regressions is expressed according to Eq. 3.3.

$$R_{it} = \gamma_{0t} + \gamma_{1t}\beta_{it} + \gamma_{2t} \ln MV_{it} + \gamma_{3t} \ln BM_{it} + u_{it} \tag{3.3}$$

Bodie, Kane, and Marcus (2014) summarizes their work and points out that higher book-to-market indicates that firms are in financial distress and that firms with lower market capitalization are more likely sensitive to business cycles. Market β in this model is still relevant as they believe that it captures the macroeconomic factors (Fama and French, 1996). Conclusively, the Fama French Three Factor Model assumes market, size, and BM to be systematic factors that constitute the conventional risk-return framework.

3.2 Credit Risk Model: The Merton Model

The Merton Model allows for the use of equity prices in order to estimate the credit risk of publicly listed companies through predicting the probability of default (Hull, 2018). The economic principles behind the Merton Model is based on the underlying characteristics of a company with corporate debt, incorporating both business risk and financial risk. These main inputs are the total market value of assets, asset volatility, and leverage, together making the structural part of the model. The larger the asset value is, all else equal, the lower the probability of default will be. The larger the asset volatility, referring to the business risk of a company, the higher the probability of default will be. Similarly, the higher the leverage is, i.e. an increased financial risk, the higher the probability of the default will be.

Through applying a somewhat simplistic view on the company value of equity as a call option on the company's assets, Merton (1974) apply the principles of option pricing from Black and Scholes (1973). The economic principle of default occurring once the asset value falls below the value of the debt at the time of the maturity is a central assumption from option pricing theory in the model, since this is when the company does not have enough funds to repay the debt-holder. Merton (1974) makes assumptions that a company cannot default before the time of maturity. The model also assumes a simplistic view of debt, as a zero-coupon bond. Furthermore, the model assumes that information relevant for probability of default is incorporated in the share price, since the price also includes the future expectations on the company. It is therefore a forward-looking structural credit risk model.

Because the model is forward-looking, the credit risk, as defined by the Merton Model, will be quicker to adjust to changes that deteriorates the credit quality of a company (Hull, 2018). The market value of equity is plainly observable through the share price of listed companies but is also only one part of the market value of a company's assets, the other being the market value of corporate debt. Both the market value and the volatility of the assets varies over time, which allows for the probability of default to also vary over time. The option pricing principle from Black and Scholes (1973) assumes the value of an asset at maturity to be log-normally distributed. The asset value at maturity and the nominal value of corporate debt is standardized so that probability of default, or credit quality, in the Merton Model becomes comparable between companies.

4. Data and Method

This chapter describes in detail which data is used in our study, the selection processes and motivation behind the quality metrics used in this study. This followed by a detailed specification of the two methods used to study our aim of this paper.

4.1 Data

The sample selection in this paper includes in total 184 active Swedish firms listed on the Stockholm Stock Exchange during the period 2010-04-01 to 2020-01-01. The data selection intends to include a longer period to reduce the variance of the coefficients by increasing the number of observations, however, due to limited access of the chosen credit risk models the period cannot be extended further. A cross-sectional regression approach is utilized in this study, thereby only firms that are traded for the whole sample period and with accessible accounting data was included to balance the sample.

Monthly Stock Prices, Dividend Payouts, Market Capitalization (expressed in million SEK), ROA, Debt to Equity (leverage) ratios, and Book Value of Equity per share are collected using Thomson Reuters/DataStream. The prices are adjusted for buybacks, splits and new issuances. After filtering for active constituents of all sizes that are primarily quoted on the Stockholm Stock Exchange and filtering out minor stocks for firms with both A- and B-shares, the original sample are 264 firms. Firms that have missing data in the beginning or middle of the sample period is excluded due to interpolation may lead to misguiding results. Some firms have missing accounting data for the whole or part of the last year due to prolonged reporting or non-calendar fiscal years. In those cases, the latest observed values are assumed. The choice of including the firms with recent missing accounting data is reasoned by the amount of observations that would be lost otherwise and since these are lagged variables anyways. Since balance sheet data are reported yearly, DataStream assumes the yearly values to be constant for all months so that it matches the rest of the data set. Monthly data is used to deal with potential problems such as downward bias that can arise if stocks are rarely traded and illiquid which could result in many zero-return observations on daily basis (Koller et al., 2015). Firms that are listed on secondary marketplaces such as First North, NGM Nordic MTF and Spotlight are excluded due to less availability of this study's quality variables. After all adjustments are made, we have 184 firms in the sample of this study.

As a proxy for the market portfolio the OMX Stockholm All-Share Gross Index (including dividend) is used and downloaded using DataStream. This index is used since excess return for the firms are including dividend (using Eq. 4.1 where r_f is the risk-free rate) and the mentioned index is broad and reflects all industries across the Swedish market. From Riksbanken (2020) the interest rate for Swedish treasury bill with three months maturity is collected and used as a proxy for the risk-free rate. A short treasury bill rate is chosen as the Swedish government is deemed as free of default risk and shorter maturity is deemed to be safer on the credit market.

$$R_{i,t} = \frac{P_t + Dividend_t - P_{t-1}}{P_{t-1}} - r_f \tag{4.1}$$

As a control variable based on Fama and French (1992), book-to-market is used by taking the monthly book value of equity per share divided by the share price (BM). The economic reason to use BM is argued by Fama and French (1992) that the risk is captured by reflecting future prospects of the firm. Those with low stock price and high book value of equity (high BM) are poor performer and therefore is punished by higher required rate of return. We choose to include this in our data set, as discussed in Section 3.1, as it has been proven and not to our knowledge challenged, that it can explain cross-sectional returns. Furthermore, Bodie, Kane, and Marcus (2014) explains that high BM firms can be interpreted as value firms meaning that much of their values are supported by their current assets while low BM called growth stocks because their valued are based on future prospects. BM is included as a control variable, not a quality metric, thereby following the Fama-French Three Factor Model presented in Section 3.1. Like most previous research, we include BM alongside firm size to make sure that size do not capture or is distorted by information that is attributable to book-to-market.

4.1.1 The Four Quality Variables

ROA is reported yearly, and we therefore assume it to be constant over the year. The ratio is used because it is a proxy for quality through the reflection of profitability. Asness, Frazzini, and Pedersen (2019) argues that if a quality measures is able to predict stock returns and is persistent, the market will price such a characteristic rationally. They show that profitability and ROA fulfil the requirement. Hagel III, Seely Brown, and Davison (2010) motivates the use of ROA as a profitability measure as it is a bottom-line measure in relation to the asset that support the firm's business. It tells if a firm is able to generate high enough sales on the employed asset and create a positive return through positive margins. Therefore, ROA is a product of both capital- and cost-efficiency of the firm.

Leverage is defined as debt over equity and reported values are quoted in percentage and yearly but assumed to be constant throughout the year. Leverage is used as a proxy for quality, on the basis of the finding of George and Hwang (2010) and Penman, Richardson, and Tuna (2007), who finds a negative relation between leverage and returns. Low leverage could be a product of high quality firm being able to self-finance their operation while high leverage firm could be a product of poor historical performance diminishing the equity value on the balance sheet. All else equal, firms with higher leverage are potentially more prone to financial distress, therefore low leverage is considered as high quality. While ROA reflects the efficiency of the firms, leverage reflects the financial structure of the firm and therefore they are complements to each other.

The Structural Credit Risk Model (henceforth, Structural model) is an expansion of the Merton model, presented in Section 3.2. The data is a daily implied letter rating available through Eikon. The model estimates the default risk on the premise that equity can be seen as a call option on the firm's assets with a one-year horizon. According to Reuters (2013) their models is superior to the Big Three credit rating agencies' models, the basic Merton model and other conventional risk models such as Altman Z-Score with an 85% success-rate to predict defaults within one year on the bottom quintile scored companies. Just as the Merton model, the Structural model uses the firm's historical leverage, asset drift and volatility of market value of the firm's assets. The output is a single probability of default that covers the specific firm and not a specific debt issuance. Reuters (2013) states that the enhanced power of their model is achieved by advanced internal modelling to determine the parameters in the Merton Model (Refinitiv, 2012). Furthermore, they adjust the definitions of the point of default and the volatility by treating balance sheet liabilities based on industry specifics. The adjustments deal with the somewhat simplistic view of debt in the Merton model. The mapping is based on historical distribution of ratings when converting the probability of defaults to letter rating, hence, mitigating some of the potential problems when assuming normal distribution in the Merton model.

The SmartRatios Credit Risk Model (henceforth, SmartRatio model), updated daily, and accessed through Eikon is an implied letter rating. The model utilizes various financial ratios to predict and assess firms' financial health and credit situation. The SmartRatios compiles industry-specific metrics and financial ratios to five main components; profitability, liquidity, leverage, coverage and growth, which are then combined into a regression framework (Reuters, 2013). Their modeling also incorporates accounting information of reported and internal modelling that adjust analyst estimates of future prospects. Furthermore, SmartRatio model assign weights to the different metrics based on what financial ratios that are most informative and important in a given sector. According to Reuters (2013), the above drivers of the model makes the SmartRatio superior to conventional accounting-based credit risk models on predicting default. Some of the ratios that are inputs to their model are, for example, earnings/tangible capital, net profit margin, EBIT/interest, debt/equity, non-performing loans (for banking industry), passenger load factor (for airlines), quick ratio, ROE growth etc. All included ratios together with the adjustment made in the modeling makes the SmartRatio model a robust proxy for quality of the firms. According to Reuters (2013), equity investment based on their risk models generates abnormal returns and there is a predictive power of future agency rating changes in their model.

Due to limitation of credit ratings for smaller firms, and the slow and inconsistent update of credit ratings from the major credit rating firms, Structural and SmartRatio are chosen proxies for quality. The ratings consist of daily updated probability of defaults translated to letter rating. Daily updates of the ratings make them a dynamic proxy for quality and an up-to-date predictor. Even though both measures are credit ratings, we include both of them as they are complements to each other and based on two distinctively different foundations of modeling and information.

Both of the daily credit ratings are quantified by assigning numerical values to the credit ratings as described by Brooks (2014). The lowest credit rating, CC, is assigned the number 0, and the highest credit rating, AAA, is assigned to the number 19, all ratings are assigned integers between 0 and 19 (for conversion clarification see Table A.1). Furthermore, the credit ratings for the 1st of each month are stated as the average including four prior daily credit ratings. The five-day average is chosen for the deviation of a single daily rating not to heavily influence the quality proxy, but also, not to ignore the variability of the daily ratings.

Since only firms that trades the whole sample period is included there is a risk of the study being influenced by survivorship bias due to some of the excluded firms are delisted. In most cases the delisted firms are such due to poor performance and low quality which leads to the included data are positive biased, however, some firms are delisted because they have been bought out of the stock exchange by a strategic or financial buyer which could indicate good performance and high quality. The main issue, however, is the lack of access to the chosen credit risk model for the delisted firm. For this reason, and the scope of the study, no further adjustment has been made to deal with the survivorship bias. Summarized descriptive statistics of the data used in this paper is presented in Table 4.1.

Table 4.1: Descriptive statistics

This Table reports the average, median, and standard deviation values, of the 184 firms over the period 2010-04-01 to 2020-01-01, based on a total of 21 528 observations for each variable.

	Size (MSEK)	BM	ROA(%)	Leverage(%)	SmartRatio	Structural
Average	18 330	0.71	4.38	70	11	12
Median	2 002	0.48	6.28	44	12	12
Std. dev.	46 795	1.19	17.52	212	3	3

4.2 Method

This section describes our two methods. First, we describe in detail how the double sorting is performed to investigate the existence of the size premium in a size-based investment strategy. Second, we describe the general Fama-MacBeth two-step approach and in detail how we apply it to examine the risk factors' explanatory power on returns across firms.

4.2.1 SMB Portfolio Construction

In order to initially evaluate the existence of size premium, that there is a positive difference between small and big stocks, we begin by creating three portfolios based on market capitalization, without controlling for quality. The firms are categorized using 30^{th} and 70^{th} percentile to form the small and big size portfolios, reclassified in the beginning of each month. Asness et al. (2018) sort the firm size into five portfolios, making the portfolio selection of small and big firms being the 20^{th} and 80^{th} percentile respectively. However, to ensure that we have enough firms in each of the size categories we choose to create three portfolios (instead of five), with slightly more generous percentiles. We choose not to use the median as Fama and French (1993) to avoid having mid-cap firms in both categories that could distort the difference between the notably small and big firms. By using more concentrated size groups, if the SMB portfolio has a positive return, the result is in fact a size effect. We compute both equally- and value-weighted monthly portfolio returns of which we then create a portfolio of excess returns, SMB, that is long in small and short in big firms. The historical expected size premium is the average of those monthly SMB returns that are tested using t-statistic.

Beside the single sorted estimation of the size premium, we double sort the firms by size and quality. For each of the four quality measures, Leverage, ROA, Smart and Structural, nine portfolios are created that are small, medium and big sized, intersected with high (good), medium and low (bad) quality values. The good, medium and bad quality uses 30%, 40% and 30% percentiles respectively, at the beginning of each month. Figure A.1 show the distribution of quality within the small respectively big firms for each of the four choices of quality measurements. Both equally- and value-weighted monthly returns are computed for each portfolio. p represents the nine types of portfolios, Eq. 4.2 for the equally-weighted returns and value-weighted returns.

$$\begin{aligned} R_{pt} &= \sum_{i \epsilon p} w_{it} R_{it} \\ w_{it} &= \frac{1}{n_{p,t-1}}, \text{ for the equally-weighted} \\ w_{it} &= \frac{MV_{i,t-1}}{\sum_{j \epsilon p} MV_{j,t-1}}, \text{ for the value-weighted} \end{aligned} \tag{4.2}$$

The size premium controlled for quality is evaluated by the monthly average of three different SMB portfolios based on the three classification of quality creating high-quality SMB, medium-quality SMB, and low-quality SMB. Note that high quality defined by leverage is measured by the lower 30% while for the other quality measures, the higher value indicates better quality.

Employing this approach, factor mimicking portfolio, has the advantage of forming portfolios based on some risk factor reflected through some firm characteristic, size (Fama and French, 1992). As previously discussed, whether or not it is the size per se that is the factor loading, whatever the background factor, it should in this case still be reflected through the SMB portfolio. Note, however, that it is under the assumption that size is the relevant firm characteristic.

4.2.2 Fama-MacBeth (1973) Two-Step Approach

Fama and MacBeth (1973) presents a method for determining if factors in an asset-pricing model, such as CAPM, may explain the variation between expected excess returns of stocks. This approach requires two steps to be completed. First, the market beta explaining the excess returns are determined by running TSR for each stock. The second step is followed by running a cross-sectional regression (CSR) of all firms' stock returns for each month on beta and other firm characteristics, such as size, to examine their power for explaining the cross sectional differences in expected excess returns.

Using TSR to estimate the beta could potentially lead to errors-in-variable which is an underestimation of the beta and lead to overestimation of other variables coefficients. The problem lies in that beta must be estimated in the TSR ($\hat{\beta}$), therefore it includes an error term in the estimation. When then using those estimated betas in the CSR, which also includes an error term, the total estimation error will not have zero-covariance to beta. Not having zero-covariance between the betas and the error term can lead to biased estimations. The errors-in-variable problem can be solved by making portfolios of the stocks in the sample, since the estimated $\hat{\beta}$ for a portfolio is closer to the true β than the estimates $\hat{\beta}$ for a single stock. According to Fama and MacBeth (1973), making portfolios of the sample reduces the variance of the error term by the diversification effect, and thereby allows for a more efficient coefficient. Litzenberger and Ramaswamy (1979) states that making portfolios of the stocks does reduce the bias in the estimated coefficients of an OLS, but the downside is that making portfolios of a sample results in the OLS coefficients being less efficient since the grouping also results in an information loss. Fama and MacBeth (1973) suggest testing the statistical significance of the average month-by-month coefficients from the CSR with the t-statistic, since the distribution of stock returns typically have higher kurtosis than a normal distribution. Since we are limited by the period length of the data at the same time as we do not want to lose information about individual firm characteristics, therefore no action is taken to deal with the errors-in-variable problem.

We perform the Fama and MacBeth (1973) two-step approach with rolling window. Since true market beta is unknown, we begin my estimating it. Fama and French (1992) estimated market beta, β_i, by creating portfolios based on size and previously estimated beta to reduce the errors-in-variable problem. Since we are, in this study, looking at firm characteristics as explanatory variable for expected stock return, and those data are individual, we estimate the firms' betas individually to not forego any information about risk factors and expected returns.

To estimate β_i for period t, we run a TSR using OLS of each stock excess returns, R_i, on market excess return, R_M, with the estimation period of the previous 36 months of observations ($s = t - 36, ..., t - 1$), instead of the conventional 5 year rolling window, given the length of the time period in our sample..

$$R_{is} = \alpha_i + \beta_i R_{Ms} + \varepsilon_{is} \tag{4.3}$$

The second step is vulnerable to the error terms of the CSR each month suffering from correlation and heteroscedasticity (Brooks, 2014). The heteroscedasticity stems from the variance of the stock returns is not constant across all stocks. Using OLS would therefore lead to inefficient estimators. One potential solution to the heteroscedasticity is to use the GLS instead of the OLS in the estimation of coefficients of the CSR (Litzenberger and Ramaswamy, 1979). However, they assume that the correlation of firms' error terms is zero in a single index model. Following Litzenberger and Ramaswamy (1979), we use a covariance-variance matrix when estimating the GLS which allows for homoscedastic error terms from the TSR.

To evaluate the cross-sectional explanation of expected returns, we first run CSR with the stocks excess return for period t on all firms ($i = 1, ..., 184$) estimated market betas up to period $t - 1$ from above (Eq. 4.3). We use the covariance matrix that allows for heteroscedasticity, but assumes zero covariance of error terms from the TSR, to estimate the GLS. This tests the explanatory power of CAPM (Eq. 4.4), if it holds γ_0 should be zero and γ_1 positive. Second, we add lagged values of log transformed market value and lagged book-to-market (Eq. 4.5), therefore testing the FF3 where our focus is the γ for $\ln MV$. Finally, we run the CSR with added lagged quality metric (Eq. 4.6). Each CSR modification is run for period $t = 37, ..., 117$, month-by-month and by moving the corresponding estimation period for β_i.

$$R_{it} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{it} + \varepsilon_t \tag{4.4}$$

$$R_{it} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{it} + \gamma_{2t} lnMV_{i,t-1} + \gamma_{3t} BM_{i,t-1} + \varepsilon_t \tag{4.5}$$

$$R_{it} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{it} + \gamma_{2t} lnMV_{i,t-1} + \gamma_{3t} BM_{i,t-1} + \gamma_{4t} Q^k_{i,t-1} + \varepsilon_t \tag{4.6}$$

Eq. 4.4, Eq. 4.5, and Eq. 4.6, use the estimated coefficients, $\hat{\beta}_i$, from Eq. 4.3. CSR on Eq. 4.6 is done separately for all four quality measures (k). From each of the six month-by-month CSRs, the coefficients are saved and tested using the t-statistic. The t-statistic determines if each of the coefficients from the CSR are statistically significant in explaining excess returns across firms, i.e. if the risk factors create any risk premium.

$$\begin{aligned} t\text{-}stat_j &= \frac{\overline{\gamma}_j}{\sqrt{\sigma^2(\hat{\gamma}_j)}} \\ \overline{\gamma}_j &= \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_{j,t} \\ \sigma^2(\hat{\gamma}_j) &= \frac{1}{T}\frac{1}{T-1}\sum_{t=1}^{T}(\hat{\gamma}_{j,t}-\overline{\gamma}_j)^2 \end{aligned} \tag{4.7}$$

The t-stat is calculated by Eq. 4.7, where $t_j \sim$ T-distribution with $T-1$ degrees of freedom. This approach uses factor loadings to determine if expected stock returns can be explained. By introducing market loading, as opposed to the method in Section 4.2.1, we also test whether CAPM holds or not, also by controlling for size and BM, and then quality we can assess if the pricing model is misspecified and that proposed the risk factors can explain the returns across stocks.

5. Results and Analysis

The first section in this chapter presents the empirical findings from the two described methods with brief comments and analyses of the results with the foundation in the previous research, theoretical models and potential pitfalls. First, how the portfolios based on size and quality performed with a focus on analyzing the results from the SMB portfolios. The second part analyses if discussed factors can describe expected returns across firms. The section is then concluded with potential pitfalls of the sample data and method.

5.1 Portfolio Monthly Average Returns

Table 5.1 presents the monthly average excess return of the single sorted portfolios based on size for the entire sample period. The average excess returns within each of the size portfolios, small, medium, and big, does have significant positive average excess returns, both for the equally- and value-weighted portfolios. However, the difference between the small and big portfolios average monthly excess returns is not significantly different from zero, in neither equally- nor value-weighted SMB portfolio, i.e. there is no significant size effect before the quality-control.

This initial result of zero risk premium of the SMB portfolios indicates that the size effect was non-existent in the Swedish Stock Market between 2010 and 2019, it follows the critique that has been proposed in the literature. In its purest form, as Berk (1995) argued, the size effect should be present since smaller firm represents more risk and should therefore yield a risk premium, though the difference was not sufficiently high enough to prove the theory. This means that either the theory does not hold, aligned with Black (1993) who argued that the lack of theory is the reason for no findings of any historical pattern, or that there are factors that distorts the size premium. Therefore, some control needs to take place to improve the size premium. In this study, it is the quality-control. Since there are arguments for higher quality of firms leading to higher stock returns, SMB entails going long in partly bad firms and short in partly good firms. The simple SMB therefore distorts the size premium when not considering the intersections of size and quality, this is evident in some of the double sorting results.

Table 5.1: Single Sorted Portfolio by Size

This Table reports the average of 117 monthly excess returns (in %), of 184 firms divided into 30/40/30 percentiles from 2010-05-01 to 2020-01-01, of the single sorted portfolio as described in Section 4.2.1. Standard error is reported in parenthesis. The *, **, *** denotes significance at the 5%, 1% and 0.1% levels respectively based on the standard t-test. All calculations are made in MATLAB.

Portfolio	***Equally-Weighted (%)***	***Value-Weighted (%)***
Big	1.34***	0.02**
	(0.0039)	(0.0001)
Medium	1.58***	0.02***
	(0.0042)	(0.0001)
Small	1.16**	0.02**
	(0.0040)	(0.0001)
SMB	-0.18	0.004
	(0.0034)	(0.0001)

Double-sorting using the ROA as the quality-control, show significant size effect through average monthly excess returns of the SMB portfolio, as presented in Table 5.2. This holds for both equally- and value-weighted, except for equally-weighted SMB with medium quality. SMB portfolios of the low-quality firms are although negative. SMB with equally-weighted returns gives 229 bps more in average monthly excess return compared to value-weighted SMB within good firms (2.53% compared to 0.24%), however, 105 bps lower within the bad firms (-1.22% compared to -0.17%). The equally-weighted SMB portfolios have higher standard error than the value-weighted portfolios. However, good equally-weighted SMB has higher standard error than the bad equally-weighted SMB (0.0078 compared to 0.0048), while as for the value-weighted construction, good has lower risk than the bad portfolio (0.0004 compared to 0.0009). Note that the size premium is larger, but also has a higher standard error compared to before the quality-control.

Table 5.2: Double Sorted Portfolios by Size and **ROA**

This Table reports the average of 117 monthly excess returns (in %), of 184 firms from 2010-05-01 to 2020-01-01, of the double sorted portfolio, sorted on size and the ROA as the quality metric as described in Section 4.2.1 using 30/40/30 percentiles. Standard error is reported in parenthesis below the average monthly excess return. The *, **, *** denotes significance at the 5%, 1% and 0.1% levels respectively based on the standard t-test. All calculations are made in MATLAB.

	Equally-Weighted (%)			***Value-Weighted (%)***		
Portfolio	**Good**	**Medium**	**Bad**	**Good**	**Medium**	**Bad**
Big	1.76***	1.35***	0.80	0.07**	0.04**	0.16
	(0.0041)	(0.0039)	(0.0049)	(0.0003)	(0.0001)	(0.0009)
Medium	2.43***	1.74***	-0.09	0.09***	0.05***	0.02
	(0.0045)	(0.0043)	(0.0053)	(0.0002)	(0.0001)	(0.0004)
Small	4.29***	1.90***	-0.43	0.31***	0.12***	-0.02
	(0.0081)	(0.0044)	(0.0043)	(0.0005)	(0.0003)	(0.0001)
SMB	2.53**	0.55	-1.22*	0.24***	0.08**	-0.17*
	(0.0078)	(0.0038)	(0.0048)	(0.0004)	(0.0003)	(0.0009)

The double sorting using the Structural model as the quality-control, in Table 5.3, show a significant size effect in the value-weighted SMB portfolios within each quality group. Similar to ROA, the good firms strengthen the size effect more than medium or bad quality size effects (it is negative for bad SMB). There is however no significance of the size effect using the equally-weighted returns. The size premium of the value-weighted SMB within good firms is 0.23%, one bps less with the corresponding portfolio using ROA. The bad SMB portfolio has an average monthly excess return of -0.25%, less than all corresponding portfolios. As with ROA, the good firms give a higher significance level of the SMB returns than the bad firms. The quality-control compared to the single sorting does in this case lead to higher risk making equally-weighted SMB portfolios insignificant, however for the value-weighted portfolios, the good SMB has the standard error of 0.0007 which is lower than for the bad firms (0.0011).

Table 5.3: Double Sorted Portfolios by Size and **Structural Model**

This Table reports the average of 117 monthly excess returns (in %), of 184 firms from 2010-05-01 to 2020-01-01, of the double sorted portfolio, sorted on size and the Structural model as the quality metric as described in Section 4.2.1 using 30/40/30 percentiles. Standard error is reported in parenthesis below the average monthly excess return. The *, **, *** denotes significance at the 5%, 1% and 0.1% levels respectively based on the standard t-test. All calculations are made in MATLAB.

	Equally-Weighted (%)			***Value-Weighted (%)***		
Portfolio	**Good**	**Medium**	**Bad**	**Good**	**Medium**	**Bad**
Big	1.33***	1.36***	1.05	0.05*	0.03	0.27*
	(0.0034)	(0.0040)	(0.0055)	(0.0003)	(0.0001)	(0.0012)
Medium	1.91***	1.72***	0.90	0.09***	0.05***	0.07
	(0.0041)	(0.0042)	(0.0060)	(0.0002)	(0.0001)	(0.0005)
Small	2.07***	1.49***	0.66	0.28***	0.07***	0.02
	(0.0051)	(0.0039)	(0.0054)	(0.0008)	(0.0002)	(0.0002)
SMB	0.74	0.13	-0.31	0.23**	0.05*	-0.25*
	(0.0039)	(0.0038)	(0.0059)	(0.0007)	(0.0002)	(0.0011)

The result from the doubled sorted portfolios aligns in parts with Asness et al. (2018), adding a quality factor to divide the SMB portfolios shows the size premium within the different quality groups except for the low-quality firms. ROA and the Structural model seem to be relevant measurements of quality to create a risk premium of the size-based portfolios. Interestingly, the different construction methods, of equally- and value-weighted portfolios, indicate that the size effect is actually a size premium. In an equally-weighted portfolio, the proportion of the smallest firms in the small group are larger than in a value-weighted portfolio. In the value-weighted portfolio the larger firms' returns in the small group are more heavily weighted. Using the equally-weighted SMB with ROA as the control gives higher size premium compared to the corresponding value-weighted portfolio, indicating that it actually is a size premium for smaller firms. The same conclusion, that the size premium is most evident in the smallest firms, is also made by Crain (2011). On the other hand, the double sorted portfolios do not consistently show that the size premium exists, rather it seems that the choice of quality variable is also of importance. The next two tables gives more insight of this problem.

Leverage as a quality-control, in Table 5.4, rather show that there is only a negative size effect within bad firms when using equally-weighted returns. That portfolio had an average monthly excess return of 10 bps less (-1.32%) than the SMB of bad firms using ROA as the quality-control. There is no positive significant returns of the SMB portfolios using leverage as a quality control, but the Structural model succeed to isolate the size effect. It could also mean that it is not entirely correct to assume it as a valid proxy for quality, but rather needs to also account for other parameters that the Merton model does.

Table 5.4: Double Sorted Portfolios by Size and **Leverage**

This Table reports the average of 117 monthly excess returns (in %), of 184 firms from 2010-05-01 to 2020-01-01, of the double sorted portfolio, sorted on size and the leverage as the quality metric as described in Section 4.2.1 using 30/40/30 percentiles. Standard error is reported in parenthesis below the average monthly excess return. The *, **, *** denotes significance at the 5%, 1% and 0.1% levels respectively based on the standard t-test. All calculations are made in MATLAB.

	Equally-Weighted (%)			***Value-Weighted (%)***		
Portfolio	**Good**	**Medium**	**Bad**	**Good**	**Medium**	**Bad**
Big	1.73***	1.29**	1.33**	0.15*	0.04**	0.05*
	(0.0044)	(0.0041)	(0.0040)	(0.0006)	(0.0001)	(0.0002)
Medium	1.85***	1.58***	1.27**	0.07***	0.07***	0.06**
	(0.0047)	(0.0043)	(0.0046)	(0.0002)	(0.0002)	(0.0002)
Small	2.00***	1.18*	0.01	0.10***	0.05*	-0.01
	(0.0047)	(0.0058)	(0.0052)	(0.0002)	(0.0002)	(0.0004)
SMB	0.27	-0.11	-1.32**	-0.04	0.01	-0.06
	(0.0048)	(0.0056)	(0.0045)	(0.0006)	(0.0002)	(0.0003)

Table 5.5 presents the SMB results using SmartRatio, it shows similar results as leverage with regards to significance of the size effect. Only on SMB portfolio show significant average monthly excess returns, although negative, in this case, the value-weighted bad firms when controlling for the SmartRatio rating. The return is one bps more negative (-0.18%) than the corresponding portfolio using ROA as the control.

The SmartRatio model and leverage of firms does not show an existence of the size premium. There are at least two possible explanations for these results. First, the SMB return was only significant for the bad firms' portfolio, and even more puzzling resulted in a negative return. This could mean that the higher quality using these measures was not necessarily good for explaining returns, especially not for the small companies. Second, the explanation for these results could be that there might not be a relationship between these two quality measurements and the size premium, at least in this sample. The leverage itself can explain firm returns, according to Penman, Richardson, and Tuna (2007), and George and Hwang (2010), but the relationship to the size premium is not as certain.

Table 5.5: Double Sorted Portfolios by Size and **SmartRatio Model**

This Table reports the average of 117 monthly excess returns (in %), of 184 firms from 2010-05-01 to 2020-01-01, of the double sorted portfolio, sorted on size and the SmartRatio model as the quality metric as described in Section 4.2.1 using 30/40/30 percentiles. Standard error is reported in parenthesis below the average monthly excess return. The *, **, *** denotes significance at the 5%, 1% and 0.1% levels respectively based on the standard t-test. All calculations are made in MATLAB.

	Equally-Weighted (%)			***Value-Weighted (%)***		
Portfolio	**Good**	**Medium**	**Bad**	**Good**	**Medium**	**Bad**
Big	1.05*	1.37***	1.64***	0.04	0.03**	0.20**
	(0.0042)	(0.0040)	(0.0046)	(0.0009)	(0.0001)	(0.0006)
Medium	1.65***	1.57***	1.40**	0.06**	0.04**	0.16***
	(0.0046)	(0.0043)	(0.0053)	(0.0002)	(0.0001)	(0.0004)
Small	1.70***	1.16*	0.76	0.16***	0.06**	0.02
	(0.0041)	(0.0044)	(0.0056)	(0.0004)	(0.0002)	(0.0003)
SMB	0.64	-0.22	-0.88	0.11	0.03	-0.18**
	(0.0043)	(0.0039)	(0.0057)	(0.0008)	(0.0002)	(0.0005)

For all quality metrics, using bad firms to construct the size-based portfolios resulted in no positive average monthly return, whereas SMB of good firms are all zero or positive. This could be explained by, in the group of small firms, the bad portion is over-represented, and the size distribution is opposite among the big firms (see Appendix A.1). This means that when constructing a SMB portfolio, without considering the quality, we go long in too many low performing firms, hence no size effect emerges. This supports the notion that it is highly relevant to include a quality metric in the decision for the construction of a size-based investment strategy.

The portfolios of small and big stocks also have relatively different distributions of quality, within the small and big groups, depending on which quality metric is used. The portfolios of large companies only have a small share of good quality stocks using the leverage and the SmartRatio model as a definition of quality, whereas the share of good quality large stocks using ROA and the Structural is greater (see Appendix A.1). Given a larger sample, these distributions of quality for leverage and SmartRatio model is likely not an issue, however, the intersection of size and quality could potentially result in too few observations for the SMB portfolio in our data sample.

The standard error for every SMB portfolios with quality-control has increased compared to the single sorting. It may seem like the increase contradicts the notion that quality helps to reduce risk. However, since we are reducing the number of firms in each portfolio due to the quality-control, we are reducing the diversification effect in order to enhance the risk premium of size and make it significant from zero.

5.2 Cross-Section of Expected Returns

The Fama-MacBeth two-step method compared to the results of double sorting evaluates the explanatory power of risk factors and other variables on expected excess returns for the individual firms. The results indicate what risk factors that are relevant to include in the investment decision and what variables to consider when forming a portfolio. The average of the coefficients, γ_i, can be interpreted as risk premium. The different specification we use in the CSR include different characteristics that are based on asset pricing theories where the factors loading of firms to these factors create a risk premium. Hence, the CSR tests if the asset pricing models are valid or not. If rejected, the reason can be that the asset model is misspecified. The results from the Fama-MacBeth regressions tells an investor which factors loading creates positive returns and are relevant to consider in the investment decision.

Table 5.6 reports the statistics using Eq. 4.7 from the 6 different Fama and MacBeth (1973) two-step regressions. The results from the CSR using the CAPM-specification is aligned with Fama and French (1992) who find no explanatory power of market beta but positive abnormal return (γ_0), meaning that CAPM does not hold or is misspecified for the period between May 2013 to December 2019. The Fama French Three Factor model (FF3) was also disproved with only γ_0 being significant, with higher average (193 bps) but at a lower significance level compared to CAPM. Surprisingly, also the explanatory power of the book-to-market was insignificant, which has not been heavily questioned to our knowledge by the literature. Before quality-control there is no size premium, consistent with the result from the single sorting and previous evidence.

Adding ROA as a quality controlling variable to the FF3 returns similar results for γ_0, both with respect to the average (at 191 bps) and the significance level. Beta and book-to-market show no significant explanation for the expected returns across firms. ROA itself has a significant 6 bps average in explaining the cross-sectional returns, indicating that higher quality yields higher returns. Most noteworthy is that $\ln MV$ gives significant negative 14 bps return, indicating that a positive size effect emerges when controlling for quality.

The size effect does not appear with the control for firm's leverage. Beta and book-to-market is also not able to explain returns across firms. γ_0 is at the same level as FF3 and ROA specification and equally significant. The leverage as a factor does significantly explain firm returns with a premium of -0.2 bps. Since it is negative, it means that higher leverage, defined as bad in terms of quality, gives lower return. The SmartRatio model as a quality-control variable result in no variable being able to significantly explain the returns across firms. Comparing to the FF3 without quality-control, the significance of γ_0 average disappears, and both book-to-market and size remains insignificant. The quality variable itself, SmartRatio, is not significantly explaining the returns across firms either, meaning it lacks explanatory power of return.

The quality-control variable that we call Structural model, does not show any significance in γ_0, beta remains insignificant different from zero, however, book-to-market is significant and the size premium reappears with -14 bps, the same as when using ROA as quality-control. The rating has a significant explanatory power with 17 bps. The standard error for beta, and size is roughly the same through all specification. Leverage variable has the lowest standard error, followed by, ROA, Structural and SmartRatio.

The results from the Fama-MacBeth regressions validates the results from the portfolio analysis. ROA and Structural model as quality variables are the only quality metrics that creates a negative size coefficient (positive size effect). The quality variables that significantly explain returns are; ROA, Structural model, and leverage. ROA has less presence in the literature as a quality variable explaining returns, but the Structural model is a Merton based model of a sound theoretical background. The theoretical background of Structural model could be the reason why it has almost 3 times higher coefficient than ROA. The leverage has substantially lower significance compared to the two others, and there is no size premium emerging from adding leverage to the specification. This means that, for a quality variable to be relevant for the size premium, it needs to have high explanatory power of expected returns across firms. That gives support for the theory that higher quality means higher returns, which contradicts the risk-return framework, and therefore quality distorts the size effect.

Table 5.6: Average and t-test of Fama-MacBeth two-step Regressions Results

This table reports the averages ($\overline{\gamma_i}$) and the standard errors of the 81 monthly coefficients of the CSR for the 184 firms, for each of the specifications explained in Section 4.2.2. The total data period for the two-step approach is 2010-04-01 to 2020-01-01. The standard error is reported in parenthesis below each $\overline{\gamma_i}$. The *, **, *** denotes significance at the 5%, 1% and 0.1% levels respectively based on the standard t-test. All calculations are made in MATLAB.

Specification	**Intercept**	β	**BM**	**size** ($\ln MV$)	**Quality**^{k}
	$\overline{\gamma_0}$	$\overline{\gamma_1}$	$\overline{\gamma_2}$	$\overline{\gamma_3}$	$\overline{\gamma_4}$
CAPM	0.0175***	-0.0010			
	(0.0036)	(0.0027)			
FF3	0.0193**	0.0020	0.0057	-0.0008	
	(0.0073)	(0.0026)	(0.0034)	(0.0006)	
FF3 & ROA	0.0191*	0.0023	0.0061	-0.0014*	0.0006***
	(0.0073)	(0.0026)	(0.0034)	(0.0006)	(0.0001)
FF3 & Leverage	0.0192*	0.0017	0.0063	-0.0007	-0.00002*
	(0.0073)	(0.0026)	(0.0034)	(0.0006)	(0.0000)
FF3 & SmartRatio	0.0106	0.0019	0.0059	-0.0008	0.0006
	(0.0104)	(0.0026)	(0.0035)	(0.0006)	(0.0004)
FF3 & Structural	-0.0013	0.0040	0.0070*	-0.0014*	0.0017***
	(0.0087)	(0.0026)	(0.0034)	(0.0006)	(0.0003)

Unlike Asness et al. (2018), it is noteworthy that in our sample it matters which quality metric is used to improve performance of a size-based investment strategy. The results indicate that ROA or Structural model rating can improve the performance when considering the size-based investment strategy, whereas leverage and SmartRatio cannot.

One explanation of why leverage failed to create any size premium could be that leverage can be seen as a risk factor that creates risk premium, rather than a quality metric as defined in this study. Penman, Richardson, and Tuna (2007), and George and Hwang (2010) find evidence of the opposite of the risk-return framework and consider it as a equity puzzle. Potentially it could mean our results confirms the notion that size effect is only a product of data mining, rather than an established theory. On the other hand, it cannot be disqualified as a valid quality metric since the relation between leverage and return is negative, meaning higher leverage (higher risk) gives less return. The SmartRatio failed to be considered as a valid quality metrics since it lacks any explanatory power of returns. Most likely that means that Reuters (2013) misspecify their models to create their ratings or is mispriced by the market. According to them, the rating helps investors to create abnormal return. Another reason could be the 5-day average used gives a wrong view of the actual quality. However, since the accounting ratios of firms only updates quarterly or annually, while the rating updates daily, this should not be a problem.

6. Conclusion and Further Research

In this paper we investigate if the size effect exists in the Swedish stock market, and if a size-based investment strategy can be improved by controlling for firms' quality. We find that there is a risk premium for size, but it is distorted by the quality of firms. The investor can improve the performance of a size-based investment strategy by controlling for the firms ROA or Structural credit rating. Otherwise, the investor risks going long in bad, low performing, firms and short in good, high performing, firms. We show this by constructing SMB portfolios that are first only sorted by size and then double sorted using the intersection of size and quality. The result is validated when performing the Fama and MacBeth (1973) two-step regressions that determines the factors relevant to explain the cross-section of returns.

The notion of size effect has been heavily questioned and disproved. Our findings partly supports the critiques, from Black (1993), Horowitz, Loughran, and Savin (2000), Gompers and Metrick (2001), Israel and Moskowitz (2013), among others, as we do not find the size premium in plain sight in the Swedish Stock Market, at least during the period 2010 to 2019. However, our findings also partially supports Asness et al. (2018). When certain quality variables are controlled for, the size effect appears. By using ROA or the Structural credit risk model, an investor can improve the performance of a size-based investment strategy by going long in firms with high quality values. On the other hand, we do not find support for Asness et al. (2018) conclusion that any quality metric isolates the size effect. Unlike the findings of George and Hwang (2010), and Penman, Richardson, and Tuna (2007), leverage failed to improve the size-based investment strategy.

By the Fama and MacBeth (1973) two-step regression approach we show that size of the firm creates a risk premium, when controlling for ROA or Structural credit rating, on monthly average -0.14% across firms. The evidence is a negative size coefficient, meaning positive size effect, i.e. smaller firms gives higher return in form of risk premium. In the CSR we assumed that market beta for individual firms are estimated by the CAPM. The evidence shows that CAPM and the Fama-French Three Factor Model does not hold in our sample data. The double sorted portfolios, using the intersection of firm size and quality show that a zero-investment

SMB portfolio creates a positive size effect using the top 30% quality firms. In our sample data, using equally-weighting and ROA creates the highest size premium of 2.53% per month on average. Whereas, using the Structural Model as the quality metric, the value-weighted SMB portfolio of the good firms created an average size premium of 0.23%. By any quality metric, creating an SMB of lowest 30% quality firms creates a negative size premium, meaning that of all bad firms, smaller firms do not outperform larger firms. For the low-quality firms, to yield positive return, the size-based investment strategy would therefore be a reverse SMB. The inconsistency of the results, where the size effect is dependent on if equally- or value-weighting is used and what quality metric is used, is noted. By using the two different methods, one important finding is that both ROA and the Structural credit rating is validated to be a relevant metric for the size effect. Additionally, the risk premium for size in the Fama-MacBeth two-step approach is the same regardless if ROA or Structural Credit Rating is used.

Going forward, it would be interesting to test these results in additional stock markets, but also using a longer time period. The longer time period would be of particular interest since it would then confirm or deny the findings throughout the entire business cycle. Some literature suggest that the size effect is prominent in January, therefore it would be of interest for the Swedish Stock Market to examine the size effect in January, and with the intersection of quality. Since some of the quality variables in this study failed to represent quality and the relation to size as expected, additional quality variables would be of interest to include in the sample data. Lastly, the unexpected result of book-to-markets failure to explain expected stock returns suggest that more research is needed on that factor, as the literature tend to focus on the size factor.

References

Asness, Clifford S., Andrea Frazzini, and Lasse Heje Pedersen. 2019. "Quality Minus Junk." *Review of Accounting Studies* 24 (1): 34–112.

Asness, Clifford, Andrea Frazzini, Ronen Israel, Tobias J. Moskowitz, and Lasse H. Pedersen. 2018. "Size matters, if you control your junk." *Journal of Financial Economics* 129 (3): 479–509.

Banz, Rolf W. 1981. "The Relationship Between Return and Market Value of Common Stocks." *Journal of Financial Economics* 9 (1): 3–18.

Berk, Jonathan B. 1995. "A Critique of Size-Related Anomalies." *The Review of Financial Studies* 8 (2): 275–286.

Black, Fischer. 1993. "Beta and Return." *The Journal of Portfolio Management* 20 (1):8–18.

Black, Fischer, and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81 (3): 637.

Bodie, Zvi, Alex Kane, and Alan J. Marcus. 2014. *Investments, Global Edition.* 10th ed. Berkshire: McGraw Hill.

Brooks, Chris. 2014. *Introductory Econometrics for Finance.* 3rd ed. Cambridge: Cambridge University Press.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi. 2008. "In Search of Distress Risk." *The Journal of Finance* 63 (6): 2899.

Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong. 2014. "Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?" *Journal of Accounting and Economics* 58 (1): 41–58.

Crain, Michael. 2011. "A Literature Review of the Size Effect." *SSRN Electronic Journal* (October).

Dichev, Ilia D. 1998. "Is the Risk of Bankruptcy a Systematic Risk?" *The Journal of Finance* 53 (3): 1131–1147.

Fama, Eugene F., and James D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (3): 607–636.

Fama, Eugene, and Kenneth French. 1992. "The Cross-Section of Expected Stock Returns." *The Journal of Finance* 47 (2): 427.

———. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33 (1): 3–56.

———. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *The Journal of Finance* 51 (1): 55–84.

George, Thomas J., and Chuan-Yang Hwang. 2010. "A resolution of the distress risk and leverage puzzles in the cross section of stock returns." *Journal of Financial Economics* 96 (1): 56–79.

Gompers, Paul A., and Andrew Metrick. 2001. "Institutional Investors and Equity Prices*." *The Quarterly Journal of Economics* 116, no. 1 (February): 229–259.

Gu, Anthony Yanxiang. 2003. "The declining January effect: evidences from the U.S. equity markets." *The Quarterly Review of Economics and Finance* 43 (2): 395–404.

Hagel III, John, John Seely Brown, and Lang Davison. 2010. "The Best Way to Measure Company Performance." Accessed April 1, 2020. https://hbr.org/2010/03/the-best-way-to-measure-compan.html.

Horowitz, Joel L., Tim Loughran, and N. E. Savin. 2000. "Three Analyses of the Firm Size Premium." *Journal of Empirical Finance* 7 (2): 143–153.

Hull, John C. 2018. *Options, Futures, and Other Derivatives.* 9th ed. Essex: Pearson Education Limited.

Israel, Ronen, and Tobias J. Moskowitz. 2013. "The role of shorting, firm size, and time on market anomalies." *Journal of Financial Economics* 108 (2): 275–301.

Koller, Tim, Marc H. Goedhart, Company McKinsey, and David Wessels. 2015. *Valuation: measuring and managing the value of companies.* 6. ed. Hoboken, N.J.: Wiley.

Kyosev, Georgi, Matthias X. Hanauer, Joop Huij, and Simon Lansdorp. 2020. "Does earnings growth drive the quality premium?." *Journal of Banking and Finance* 114.

Lintner, John. 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): 13–37.

Litzenberger, Robert H., and Krishna Ramaswamy. 1979. "The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence." *Journal of Financial Economics* 7 (2): 163–195.

Merton, Robert C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *The Journal of Finance* 29 (2): 449.

Morningstar. 2020. "Fonder, Sök." Accessed May 12, 2020. https://www.morningstar.se/se/screener/fund.aspx#?filtersSelectedValue=%7B%22categoryId%22:%7B%22id%22:%22EUCA000546%22%7D%7D&sortField=returnM60&sortOrder=DESC.

Mossin, Jan. 1966. "Equilibrium in a Capital Asset Market." *Econometrica* 34 (4): 768–783.

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

Penman, Stephen H., Scott A. Richardson, and Irem Tuna. 2007. "The Book-to-Price Effect in Stock Returns: Accounting for Leverage." *Journal of Accounting Research (Wiley-Blackwell)* 45 (2): 427–467.

Refinitiv. 2012. "Starmine Quantitative Analytics." Accessed April 2, 2020. https://www.refinitiv.com/content/dam/marketing/en_us/documents/brochures/starmine-quantitative-analytics-brochure.pdf.

Reinganum, Marc R. 1983. "The anomalous stock market behavior of small firms in January: Empirical tests for tax-loss selling effects." *Journal of Financial Economics* 12 (1): 89–104.

Reuters, Thomson. 2013. "Thomson Reuters Starmine Credit Risk Models." Accessed April 1, 2020. https://economia.uniroma2.it/public/finance/files/StarMine_Credit_Risk_Models%5B1%5D.pdf.

Riksbanken. 2020. “Sök räntor & valutakurser.” Accessed April 2, 2020. https://www.riksbank.se/sv/statistik/sok-rantor--valutakurser/?g6-SETB3MBENCH=on&from=2009-12-01&to=2020-04-02&f=Month&c=cAverage&c=cUltimo&s=Dot.

Schwert, G. William. 2002. *Anomalies and Market Efficiency.* Working Paper, Working Paper Series 9277. National Bureau of Economic Research, October.

Sharpe, William F. 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*.” *The Journal of Finance* 19 (3): 425–442.

StarMine®. 2020a. “TICKER—Models - SmartRatios Credit Risk (SRCR).” Accessed April 2. https://professionalhelp.starmine.com/help/ticker/tickers.phtml?page_set=ticker_srcr.

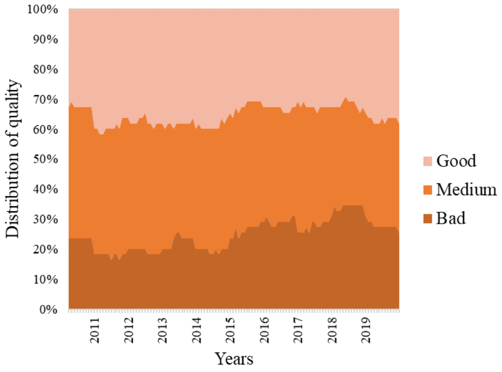
———. 2020b. “TICKER—Models - Structural Credit Risk (SCR).” Accessed April 2. https://professionalhelp.starmine.com/help/ticker/tickers.phtml?page_set=ticker_scr.

A. Appendix

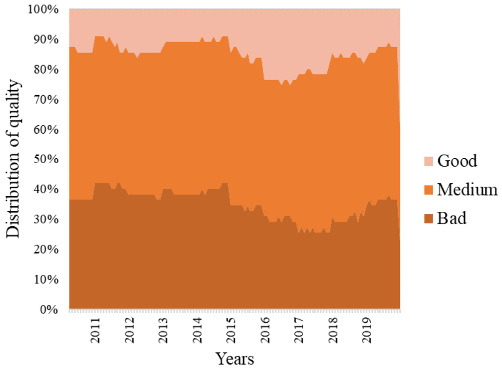
Table A.1: Thomson Reuters StarMine Credit Risk Model system

	Probability of default				
	Structural		SmartRatios		
Letter rating	Lower limit	Upper limit	Lower limit	Upper limit	**Numerical rating**
AAA	0.00%	0.00%	0.00%	0.03%	19
AA+	0.00%	0.00%	0.03%	0.04%	18
AA	0.00%	0.00%	0.04%	0.05%	17
AA-	0.00%	0.01%	0.05%	0.08%	16
A+	0.01%	0.02%	0.08%	0.11%	15
A	0.02%	0.03%	0.11%	0.16%	14
A-	0.03%	0.04%	0.16%	0.23%	13
BBB+	0.04%	0.05%	0.23%	0.32%	12
BBB	0.05%	0.07%	0.32%	0.42%	11
BBB-	0.07%	0.11%	0.42%	0.56%	10
BB+	0.11%	0.19%	0.56%	0.74%	9
BB	0.19%	0.31%	0.74%	0.99%	8
BB-	0.31%	0.47%	0.99%	1.34%	7
B+	0.47%	0.87%	1.34%	1.83%	6
B	0.87%	1.56%	1.83%	2.46%	5
B-	1.56%	2.50%	2.46%	3.36%	4
CCC+	2.50%	3.69%	3.36%	4.56%	3
CCC	3.69%	5.06%	4.56%	6.31%	2
CCC-	5.06%	7.08%	6.31%	9.66%	1
CC	7.08%	100.00%	9.66%	100.00%	0

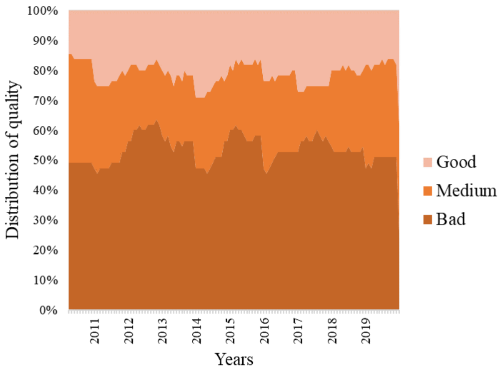
Source: Structural: StarMine® (2020b), SmartRatio: StarMine® (2020a)



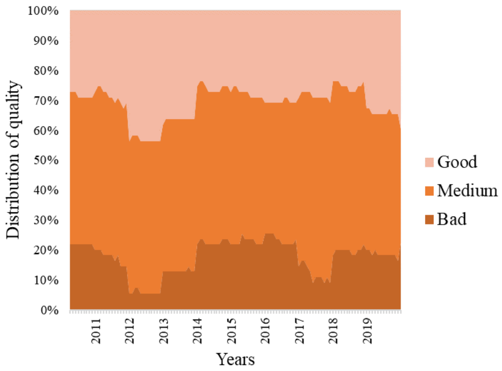
(a) Small Stocks, Leverage



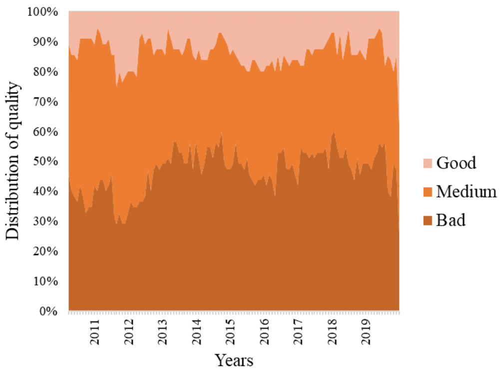
(b) Big Stocks, Leverage



(c) Small Stocks, ROA

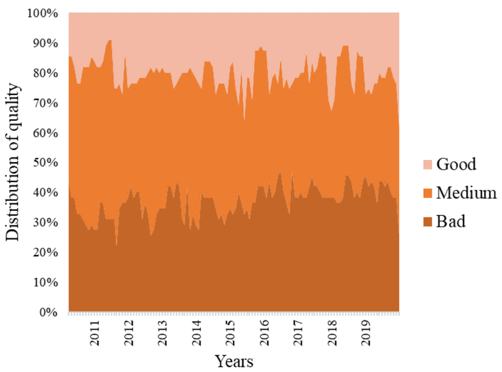


(d) Big Stocks, ROA

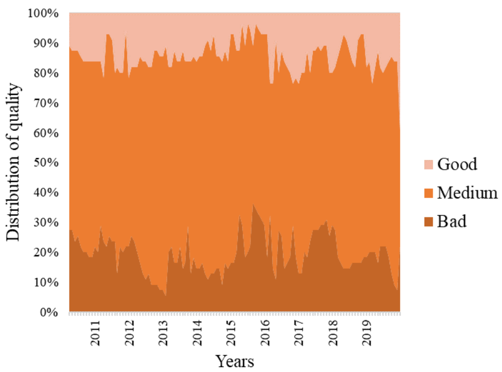


(e) Small Stocks, Structural Model

(f) Big Stocks, Structural Model



(g) Small Stocks, SmartRatio Model



(h) Big Stocks, SmartRatio Model

Figure A.1: Quality distribution in small and large stocks, where the quality percentiles are 30-40-30 (good-medium-bad), and small/big stocks are the 30% smallest/largest each month of the sample period.