ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
MSc Economics & Business
Master Specialisation Financial Economics

The impact of company name fluency and institutional ownership on stock returns and firm value

North-American common stock between 1983-2008

Author: T. Wols
Student number: 375008
Thesis supervisor: Dr. J.J.G. Lemmen
Finish date: January 2019

PREFACE AND ACKNOWLEDGEMENTS

I would like to thank Dr. M. Montone for supervising me and helping me during the first phases of my research, and Dr. J.J.G. Lemmen for being so kind to assist me during the latter phases.

ABSTRACT

I expand upon the fluency measure provided by Green & Jame (2013) and find that, across all North-American common stock between 1983-2008, company name fluency has a significantly positive effect on firm value in the cross-section. Judging from raw data and simple statistical analysis, there seems to be a positive relationship between company name fluency and stock returns. However, when taking into account risk-adjusted returns rather than real returns, these effects vanish, leading to no clear-cut relationship between company name fluency and stock returns. Therefore, the higher level of returns are likely to be a compensation for additional risk exposure. Additionally, Using two common proxies for levels of institutional ownership, I find that higher levels of institutional ownership go hand-in-hand with higher levels of risk-adjusted returns while performing portfolio analysis, and lead to higher levels of firm value in the cross-section. However, when levels of institutional ownership are high due to large concentrations of stocks held by few institutions, the relationship between institutional ownership and firm value is significantly negative.

Keywords: fluency, company name, stock returns, institutional ownership, portfolio analysis, firm value

JEL Classification: G11, G14. G23, G41

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1 Introduction

Through quantification of behaviorally economic influences, assessing their impact on stock returns is made possible. By performing portfolio analysis, I find that there is a positive relationship between stock returns and company name fluency. When looking at risk-adjusted returns however, rather than real returns, this relationship ceases to exist. Therefore, the higher levels of returns most likely serve as compensation for higher levels of risk exposure.

Itzkowitz, Itzkowitz & Rothbort (2015) state that, when individuals have to make choices involving a large number of options, they often choose the first acceptable option rather than the best available. Given this finding, I expand upon the company name fluency scores provided by Green & Jame (2013), by adding an extra criterion to the fluency score categorization. This criterion is the alphabet score, which is based on the place of the company name's starting letter in the alphabet. Whenever a company name in the sample starts with either letters A, B or C, the current fluency score effectively increases by one. This leads to a total range of company name fluency scores between zero and five, instead of zero and four.

I first perform analysis based on the old company name fluency scores, and thereafter test for the effect of fluency after the addition of the new fluency score parameter. In neither the cross-section, nor during portfolio analysis, the expansion of the fluency score (by adding the alphabet score) leads to different conclusions than before the expansion.

In the cross-section, company name fluency has a positive effect on both proxies for firm value, while controlling for multiple other, mostly macro-economic, variables. Therefore, in order to maximize firm value, company name fluency has to be taken into account. This finding speaks to the fact that a clever manager is more likely to opt into a fluent, or a more fluent, company name.

I use two different measures for institutional ownership, namely the total percentage of shares held by institutional investors and the Herfindahl-Hirschman index for institutional ownership. I find a twofold relationship between both institutional ownership and stock returns, while performing portfolio analysis, and between institutional ownership and firm value in the cross-section. First-off, I find that higher levels of institutional ownership, percentage wise, are accompanied by higher levels of both risk-adjusted stock returns and firm value. However, as the institutional holdings in a company become more clustered in very few institutions, the impact of institutional ownership on both firm value and stock returns deteriorates, and may even become negative.

Lastly, I find a negative relationship between the total percentage of institutional ownership in a company and the company name fluency in the cross-section. This means that institutional investors are less likely to invest in fluently named companies compared to non-institutional investors. In conjunction with my stated finding on the relationship between company name fluency and risk-adjusted returns, this finding could be a sign of expertise of institutional investors. The logic behind this is that due to mental shortcuts, individual investors are more likely to invest in memorable stocks, such as the ones with fluent company names.

My findings on the relationship between fluency and stock returns are not in line with most literature. For example, Head, Smith & Wilson (2009) find that company ticker fluency is positively related to stock returns. They argue that perhaps a clever ticker symbol is a useful indicator of managers' ability that revealed itself over time as the company repeatedly exceeded investors' expectations. There is however a large difference between company name fluency and company ticker fluency, which is a three to five lettered abbreviation of the company name.

On the other hand, my findings in the cross-section, on the relationship between company name fluency and stock returns, correlate well with the results found by Green & Jame (2003). The authors state that companies with short, easy to pronounce names, have higher valuation ratios. The dataset on company name fluency used in this research originates from them.

The central research question is as follows: ''Is there an effect of company-name fluency and institutional ownership on risk-adjusted stock returns and firm value?''. These are essentially four research questions which are further outlined in the data section.

The remainder of this article is subdivided into five more sections. Section two will cover the most relevant literature regarding fluency, institutional ownership, the alphabet-effect and size. Thereafter, Section three will describe all the datasets used in this research thoroughly and cover all hypotheses. Subsequently, Section four will provide an overview of the methods used in conducting the empirical research. The results of this empirical research are presented in Section five. Section six concludes the research and provides insights and recommendations for further research.

CHAPTER 2 Literature

2.1 The relevance of fluency

When the amount of information gets too large, or when decisions get too complicated, people tend to make use of decision heuristics. Heuristics, commonly defined as cognitive shortcuts or rules of thumb that simplify decision making, can be seen as rules or strategies for information processing. These heuristics help to come up with a quick, but not necessarily optimal, decision (Gigerenzer & Gaissmaier, 2011).

Huang (2005) claims that sometimes heuristics can be beneficial for making adequate decisions, while at other times, heuristics can inversely affect decision making. The reason for this mixed, or nuanced, answer, is that heuristic based actions are faster than rationally deliberated actions, which turns out to be beneficial for decision making. However, because of their impact on the speed of decision making, heuristics can lead us to make systematic errors in decision making.

Existing literature discusses the correlation between fluency and stock returns. Their results show different psychological effects and influences, which play a vital role in predicting stock prices and creating a profitable portfolio. For years, traditional finance has always presumed that, when it comes to risk-return trade-offs and maximizing utility, investors are rational in their decision-making process. However, multiple behavioral finance studies revealed that individuals do not act as rationally as economic theory assumes, due to individuals using mental-shortcuts in their decision-making process. The remainder of this sub-section touches on the psychological rationale behind fluency.

According to Adam Alter and Daniel Oppenheimer, psychologists at Princeton University, New Jersey, it's all about fluency. In trying to process complicated or large amounts of information, individuals tend to focus on the simplest parts. Therefore, in essence, individuals naturally favor things that are more fluent, which are therefore easier to think about. Hence, it is humane to try and process information in the simplest way possible (Alter & Oppenheimer, 2006, 2008, 2009).

Since individuals, and therefore investors, attempt to simplify the decision-making process as much as possible, a lot of valuable information goes to waste. In making investment decisions, individuals tend to make decisions based on heuristics, which more often than not leads to suboptimal decision outcomes (Tversky & Kahneman, 1973).

One of these non-rational heuristics used in making investment decisions, as shown by (Green & Jame, 2013), is company name fluency. Judging from normative investment theory, investors are supposed to make investment decisions based on expected return, risk, correlation with other stocks, etcetera. However, over time it becomes more evident that investors tend to invest as if companies which they are more familiar with, are better investment objectives than non-familiarized stocks. Due to a more fluent name sounding more familiar, it is therefore easier to process, leading to more investor attention (Alter & Oppenheimer, 2006, 2008, 2009).

According to Itzkowitz, Itzkowitz & Rothbort (2015), individuals are satisfiers. Typically, stock information is presented in alphabetical order, based on either the company name or the ticker, and cannot be resorted. Due to the large number of stocks in the investment universe, investors tend to pick the first acceptable option, rather than searching for the most optimal. According to the authors, early alphabet stocks have up to 22% higher turnover rates than later alphabet stocks after univariate analysis. Multivariate analysis additionally shows higher turnover, trading volume, dollar volume and lower levels of the Amihud (2002) liquidity measure.

Lastly, Ferson, Sarkissian & Simin (1999) simulate a hypothetical story in which a researcher discovers the alphabet investment strategy, which is completely unrelated to systematic risk. They find that the average value of the premium of early alphabet firms, compared to late alphabet firms, was highly significant at about one-third of a percentage point per month. However, after measuring the alpha risk factor, the regressions seemed to strongly suggest a risk-based explanation for the alphabet anomaly.

2.2 Fluency in chronological order

Numerous studies have in fact established that psychological factors do have relationships to, and impact on, the decision making of individuals. In light of this, I attempt to sketch the progress of research on the psychological factor of interest, namely fluency, over the last twelve years. The next few paragraphs will cover the most relevant research in chronological order.

According to Dreman (2004), despite what many economists and financial theorists assume, people are not good intuitive statisticians, in particular under difficult conditions. They fail to calculate odds properly when making investment decisions, which leads to persistently consistent errors. When investors are swamped by information, they (sub)consciously react only to small parts of the information. Therefore, investors usually make more suboptimal decisions than normative theories would suggest. He finds that people react to large amounts of data by adopting mental shortcuts or simple rules of thumb, instead of calculating the odds of a given outcome and making a conscious decision based on that. Psychologists call these shortcuts judgmental heuristics, which are learning and simplification strategies people use for processing large amounts of information.

Shortly after this research, Alter & Oppenheimer analyzed how well companies performed on the basis of company ticker symbols, which are self-created by companies. Their results imply that simple, cognitive approaches to modeling human behavior, sometimes outperform the typical, but more complex alternatives. For example, pronounceable symbols, such as KAR, tend to do much better than unpronounceable ones such as RDO. In one study, a pair of individuals invested their fictitious $1,000 and found that fluent codes were $85 up on the first day. even though the portfolio was just $20 ahead after a year, the fluently named portfolio still outperformed the non-fluent one.

In September of the same year, Pensa (2006) came up with supportive evidence to the finding that company name characteristics affect buy and sell decisions. Furthermore, Pensa finds that respondents to a questionnaire much rather attribute positive stock performance to a nicely sounding name than to a hideous one. Further results indicate that, not only do firms with a favorable name rating exhibit significantly higher initial one day returns after going public, but that they also generate higher risk-adjusted stock returns up to 10 trading days after the initial offering. In line with Alter & Oppenheimer (2006), these results confirm that respondents instinctively assume that firms with more fluent names have better future prospects. Additionally, they show that short-run stock performance (up to ten days after the initial public offering) can be predicted by simple, cognitive criteria, like sympathy, ease of memorization and fluency.

Next off, Head, Smith & Wilson (2007) find that stocks with memorable ticker symbols generate returns higher than average market returns. The authors propose two different explanations to support their findings. Firstly, they hypothesize that creating a clever ticker symbol is an indicator of the current manager's ability. Secondly, they hypothesize that a more fluent name increases investor attention. Their results support the first hypothesis, since the second hypothesis would most likely lead to overpricing and therefore lower returns.

A few years later, Li, Mahani & Sandhya (2011) find an effect of investor attention on short-term stock returns. They create attention portfolios, which consist of stocks whose ticker symbols are similar to stocks of large, news covered, companies (proxied by extreme returns or high trade volumes). The authors focus mainly on the effect of investor attention on prices of small-cap stocks, and find that the so-called attention portfolio, containing the stocks with the highest level of investor attention, experiences higher levels of trade activity and returns than its control portfolio in all three weeks following its formation.

More recently, Green & Jame (2013) investigate the effect of fluency on certain firm characteristics, and find that companies with more fluent and more pronounceable names have a higher breadth of ownership, greater share turnover, lower transaction costs and higher valuation ratios. They observe that individuals predict higher future returns for fictional companies with more fluent names, along with the fact that the more difficultly named company stocks are perceived to be riskier. Their findings, lastly, support the view that company name fluency has a significant effect on investor recognition and firm value.

In a very recent article, Xing, Anderson & Hu (2016) evaluate the impact of likeability and pronounceability of ticker symbols on firm value. They find that higher likeability of ticker symbols is accompanied by a higher firm value, using Tobin′s Q as a proxy. The pronounceability of ticker symbols has a similar, however weaker, effect, according to their findings. Further evidence suggests that this effect possibly exists due to the impact of company ticker symbols on stock liquidity, mispricing, or possibly both simultaneously.

Even more recently, Chan, Park & Patel (2018), find an impact of company name fluency on venture investment decisions and IPO underpricing. They apply their fluency research to renew the cognitive foundation of new venture investment decisions research. They find that both linguistic and phonetic fluency influence investment decisions.

Most recently, in June of 2018, Van den Assem, Montone & Zwinkels (2018) propose two different takes on the possible impact of company name fluency on stock returns. Firstly, higher company name fluency could attract more unsophisticated investors, which in turn leads to overpricing and subsequently lower returns. On the contrary, they hypothesize that, due to a relationship between fluency and managerial ability, returns should be higher for more fluent companies, due to failure in recognizing this information. Their empirical findings, which show higher risk-adjusted returns for more fluency companies, support the second hypothesis.

The general consensus in the aforementioned literature is that fluency, in the form of company ticker fluency, has a positive influence on stock returns. Additionally, Company name fluency appears to positively affect firm value. Not only does fluency have an influence on stock returns and firm value, it also proxies managerial capabilities, risk assessments and many other things, including lots of firm characteristic values.

2.3 Institutional ownership

Asquith, Palthak & Ritter (2005), find that constrained stocks between 1988-2002 underperform by a significant 205 basis points per month. Their measure for short-selling constraints is based on proxies for short interest and institutional ownership. Due to companies with high levels of institutional ownership being less constrained, this speaks to higher levels of institutional ownership being accompanied by higher levels of stock returns.

Blume & Keim (2014) state that the profession of investing seems to have accepted the fact that institutions tend to overweight stocks with high levels of market capitalization. However, their empirical results show that after 1990, institutions have underweighted the stocks that made up the top forty percent of market value. They show that, between 1980-2010, institutions steadily increased their holdings in smaller securities, which they define as the ones that make up less than twenty percent of total market value. They attribute these changes to the growing awareness of the size premium, which led to more investments in smaller stocks.

Gompers & Metrick (2001) denote a positive relationship between institutional ownership and future stock returns. Additionally, Gompers & Metrick find that large institutional investors nearly doubled their share of the stock market between 1980 and 1996. Due to institutions being limited to buying high-rated large stocks, this increases the relative demand for large companies, therefore decreasing the relative demand for smaller ones. According to Gompers & Metrick, this compositional shift alone can account for nearly 50% of the increase in price of large stock and part of the disappearance of the small-company premium.

In turn, Yan & Zhang (2007) show that the positive relationship shown by Gompers & Metrick (2001), between institutional ownership and future stock returns, is driven by institutions with a small investment horizon, whose goal is to achieve high returns over a short period of time. Their results are consistent with the view that short-term focussed institutions are more well informed and actively trade to exploit this informational advantage. Additionally, short-term focussed institutions’ trading forecasts future stock returns.

Edelen, Ince & Kadlec (2016) in a recent article show an exactly opposite effect of institutional ownership on stock returns. They find that institutions tend to buy overvalued stock, which leads to negative ex-post abnormal returns. They attribute this difference to their investment horizon. On a quarterly horizon, they find positive returns, however, over a one-year horizon, the results are negative.

Lewellen (2011) documents that returns and stock holdings between 1980 and 2007 show no evidence of superior stock-picking abilities for institutional investors compared to individual investors. The author claims that institutions as a whole closely mimic the market portfolio. A very interesting finding is that apparently, institutions do not seem to invest based on any of the known anomalies, such as book-to-market and momentum. Even though some institutions show higher levels of stock-picking skill relative to the Capital Asset Pricing Model (CAPM), their performance can almost entirely be attributed to book-to-market and momentum return effects, and thus risk.

Holcomb, Holmes & Connelly (2009) show that managers are a potential source of value creation for a firm. By using data from professional sports teams, they test their managerial ability theory. Even though their results show indication of the influence of managerial ability on firm value, this effect is less pronounced for increases in the quality of firm resources.

Intuitively, sophisticated investors should not be prone to making investment decisions based on heuristics and mental shortcuts. However, Coval & Moskowitz (1999) show that institutional investors do prefer investing in locally headquartered firms. This so-called home bias does not lead to additional systematic risk exposure, but since the institution, due to its proximity, is exposed to similar risks as local companies, this leads to lower levels of diversification than when non-local alternatives were to be included in the portfolio instead.

Similarly, Grullon, Kanatas & Weston (2004) find that institutional investors are prone to invest in the more familiar. Specifically, they find that institutional investors are more likely to invest in firms that advertise heavily, which means exposure to the so-called attention bias. This bias states the tendency for people’s perceptions to be affected by recurring thoughts. Higher levels of news coverage cause more recurring thoughts, and therefore speak to the attention bias.

Hung (2014) finds similar results for the relationship between institutional ownership and psychological biases, especially the attention effect. As one explanation for why institutional investors suffer from the attention bias, the author states an agency problem. In this specific case, agency problems may exist due to managers investing money based on their own interest rather than that of the individual investors that deposited money into the fund. Additionally, fund managers are not evaluated based on the performance of their individual funds, but rather on the performance of the entire fund company itself. Hung argues that this might lead to collusion, at the expense of investor returns.

CHAPTER 3 Data

3.1 The datasets explained

The main dataset consists of 1,104,861 observations across 14,856 unique companies, listed across all North-American stock markets. The data ranges from February 1983 to December 2008, which makes for a total of 299 monthly observations of fluency, factor loadings and returns. Since macro-economic data is only available on a yearly or quarterly basis, these variables are extrapolated and sometimes mutated to best match their value for each month.

Data on company name fluency, which is the psychological variable of interest in this research, is provided by Green & Jame (2003). This dataset contains yearly fluency scores of all common stock contained in the Center for Research in Security Pricing (CRSP) monthly return file and the Compustat fundamentals annual file between 1982-2009. In this dataset, depositary receipts, real estate, investment trusts and closed-end funds are not taken into account.

Green & Jame (2003) measure fluency completely objectively, therefore bypassing one of the main critiques on most behavioral research. They measure the fluency score as the sum of three independently measured proxies, which are the Englishness score, the dictionary score and the length score.

I expand upon the company name fluency scores provided by Green & Jame (2013) by adding another fluency measure. The added fluency measure to company name fluency is the alphabet score, which is based on the starting letters of each company name. After the addition of the alphabet score, fluency scores range between zero and five. Company names with a fluency scores of five being the most fluent, and company names with a score of zero being the least fluent.

Firstly, the Englishness score is derived from the frequency of the most common English letter clusters in company names. To achieve this, Green & Jame use a linguistic algorithm developed by Travers & Olivier (1978) to assess the “Englishness” of a given word. To make for even less noise, and to not favor longer company names, they focus solely on the word in the company name with the lowest Englishness score and rank all companies based on their minimum score. Subsequently, companies in the bottom quintile are given an Englishness score of zero and all other companies are given an Englishness score of one.

Secondly, the dictionary score captures whether the company name consists of words that are contained in an English dictionary. This measurement is operationalized by putting all company names through Microsoft spell check in all lower-case letters. Company names that pass the spell check take on a dictionary score of one, company names that do not pass the spell check take on dictionary scores of zero.

Thirdly, a length score is assigned to each company name. To do this, Green & Jame (2013) remove certain expressions that are part of the legal name, such as Co., Corp., Inc., Ltd., LLC, and FSB if they are the last expression in the company name. They also exclude all conjunctions in each company name, the state of incorporation that is often reported in bank names, hyphens and lastly .com. After this purification, the number of words in the remaining company name are counted. Company names consisting of one word take on a length score of two, two-worded company names take on a value of one and company names consisting of three or more words, after purification, take on a length score of zero.

Lastly, the alphabet score is based on the starting letter of each company name at any point in time. The alphabet score can take on two values, one and zero. An alphabet score of one is assigned to a company whenever a company name starts with either letters A, B or C. Whenever the company name does not start with either of these three letters, an alphabet score of zero is assigned to the company.

Data on stock returns are also retrieved from the CRSP database. Additionally, share prices and the number of shares outstanding are taken from the CRSP database in order to compute market values for each company and to proxy for company size. These two variables are further used in this research to determine book-to-market ratios and for portfolio creation.

Most macro-economic variables are retrieved from Compustat. In order to perform proper cross-sectional regressions including relevant control variables, the following variables are computed from data derived from the database, namely; sales, profitability, sales growth, asset turnover, R&D sales, advertising sales, leverage, payoutrate and Tobin’s Q.

Tobin’s Q calculations are based on Chung & Pruitt (1994), and are calculated using the following formula:

$$Tobin's\ Q_{t,i} = \frac{Total\ Assets_t - Equitybook_t + Market\ Capitalization_{t,i}}{\text{Total Assets}_t}$$

In this formula, Total Assets stands for the total average number of assets that a company has in year t. Similarly, Equitybook and Market Capitalization stand for the average book value of equity in year t and the average market capitalization for month i and year t, respectively. In the case of Total Assets and Equitybook, the average of these values is the average between their values at time t-1 and time t. For example, the average total assets in 2003 is the average between the total assets at the end of year 2002 and at the end of 2003.

Furthermore, the denotation of month i is especially important whenever data is reported on a monthly basis. This distinction is necessary since I perform regressions based on monthly data. Therefore, for the month July of 2003, Tobin's Q is based on the average total assets and the average book value of equity in 2003 (average between the end of 2002 and 2003) and the reported market capitalization in July of 2003 (t =2003, i = 7). Therefore, in essence, Tobin's Q is relatively stable for each month in a year, the only variance being the book value of equity.

I retrieve data on institutional ownership from the Thomson-Reuters database. It contains quarterly data on the two used proxies for institutional ownership. The first proxy is the Herfindahl-Hirschman index, which is highest when one institution owns all stock of a company, and decreases as either the number of institutions holding company stock increases, or when the total percentage of institutional ownership in a company decreases. The second proxy is the percentage of total common stock owned by institutions, which is a more fair proxy when distinguishing between levels of individual and institutional ownership. Since the Herfindahl-Hirschman index increases as the total percentage of stocks held by institutions increases, these two variables are closely related.

Finally, I use data, archived by Kenneth French, on factor loadings. This data is used in order to perform three- four- and five-factor model regressions. This dataset consists of monthly, American, market risk- (MRP), size- (SMB), book-to-market- (HML), momentum- (MOM), profitability- and investment-premia.

3.2 The hypotheses

The first goal of this research is to assess whether company name fluency has an effect on risk-adjusted stock returns. Based on prior research, I hypothesize the following:

H1: More fluently named company stock outperforms less fluently named company stock on the North-American stock market.

In order to formulate a complete answer to this hypothesis, it is important to make sure that fluently named companies not only generate higher returns, but more importantly that these do not serve as a compensation for risk. Therefore, I create the following two sub-hypotheses.

H1.1: There is a positive relationship between company name fluency and stock returns on the American stock market.

H1.2: The excess return generated by more fluently named stocks cannot be classified as a compensation for increased risk exposure.

The idea is that the first sub-hypothesis answers whether there is a difference in returns between fluently and less-fluently named companies. The second sub-hypothesis is in place as a check to make sure that if hypothesis 1.1 holds true, there is actual outperformance, and not just additional returns as a compensation for risk.

This is done by looking at the difference between the most fluent and the least fluent portfolio. To compare the returns of the most and least fluent companies, a zero-investment portfolio is created, going long in the most fluent and short in the least fluent category of stocks. The returns of these portfolios is then tested in a three- four- and five-factor model to assess whether the difference is due to higher risk exposure or due to actual abnormal returns achieved by the strategy.

Secondly, I am interested in the effect of company name fluency on firm value in the cross-section. I expect the same effect of company name fluency on firm value as on stock returns, and hypothesize the following:

H2: Higher levels of company name fluency lead to higher levels of firm value.

Since I am not only interested in the effect of company name fluency on stock returns and firm value, I hypothesize similarly for the effect of institutional ownership on both of these variables.

H3: Companies with higher levels of institutional ownership generate higher stock returns, disregarding transaction costs.

H4: Companies with higher levels of institutional ownership have higher firm values.

The rationale behind these hypotheses is that institutional investors are supposedly more knowledgeable, educated and devote more time to making investment decisions. Ignoring transaction costs, this should result in higher returns for institutional investors than for individual investors. Therefore, companies with higher levels of institutional ownership, are likely to generate higher levels of (risk-adjusted) stock returns. The same reasoning goes for firm value as for stock returns.

CHAPTER 4 Methods

First off, to give any incentive to do further research, the raw data needs to be looked at to see if there is any sign of a relationship between stock returns and fluency. This is done through the use of a t-test. Given that this relationship exists, more thorough research should be done to assess whether the difference in returns is due to either mispricing or risk compensation. This is done through running three- four- and five-factor model regressions based on portfolios for each one-year lagged aggregated fluency score.

Firstly, six portfolios are made, based solely on one-year lagged fluency scores. To establish whether there is any indication of a relationship between company name fluency and stock returns, I perform one-sided t-tests on the difference in returns between the least fluent and the most fluent portfolio. Additionally, I test the difference in returns between the most and least fluent portfolios before adding the alphabet score as an additional determinant for fluency score.

$$T = \frac{R_1 - R_2}{\sqrt{\frac{\sigma}{N}}}$$

The above equation describes the formula used to perform this test. Here R_1 stands for the average return on the first portfolio, and R_2 stands for the average return on the second portfolio. Based on previous literature, I expect to find here, that more fluently named companies on average earn higher returns than their less-fluent counterparts. Therefore, I expect a significantly positive t-statistic.

Consecutively, I create double-sorted portfolios (6x10) based on one-year lagged, expanded, fluency scores and size deciles. The average returns of each portfolio are regressed on multiple established risk factors, taking into account additional risk exposure for each portfolio. Therefore, portfolio returns are tested against both Fama & French three- and five-factor models and the Carhart four-factor model. (Fama & French, 1993, 2016) (Carhart, 1997).

$$R_{p.t} = \propto\ + \beta_1\, MRP_t + \beta_2\, SMB_t + \beta_3\, HML_t + D_1\, \beta_4\, MOM_t + D_2\, \beta_5\, INV_t + D_3\, \beta_5\, PROF_t + \varepsilon_t$$

The main variable of importance in this equation is $\propto$, which denotes the excess-returns generated by a portfolio. R_p stands for the average excess-return earned by a specific portfolio, and the terms followed by betas denote the return generated by zero-investment portfolios based on anomalies that are risk-driven. The corresponding betas stand for the exposure of the regressed portfolio to these risk factors, respectively loaded on the market risk premium (MRP), the size premium (SMB), book-to-market premium (HML), momentum premium (MOM), Investment premium (INV), and lastly the profitability premium (PROF).

D_1, D_2 and D_3 denote dummy variables, these are used to create a regression formula universally useable for each factor model. In testing the Fama & French 3-factor model, all three dummy variables take on a value of 0, leading to the exclusion of the momentum, investment and profitability factor. Similarly, to form a Carhart four-factor model, D_2 and D_3 take on a zero value, and to form a Fama & French five-factor model, only D_1 takes on a zero value. Lastly, the error term (ε) stands for the residual of the regressions model, which captures all variance that is uncaptured by factors included in the model. This is often referred to as the idiosyncratic risk.

In order to statistically test whether there is a difference in risk-adjusted returns between the most- and the least fluent portfolios, I create zero-investment portfolio, going long in the most- and short in the least fluent stocks. The illustrated tables in this thesis will be based on the results of Carhart four-factor regression models.

The efficient market hypothesis states that the alpha of each portfolio should equate to zero. However, what I expect to find is that the portfolios which contain the most fluent stocks, have higher risk-adjusted returns than the portfolios containing the least fluent stocks. Therefore, a zero-investment portfolio going long in the most fluent stocks and short in the least fluent stocks expectedly should have positive risk-adjusted returns. This can only hold true if the alphas generated by the most fluent portfolios are either significantly more positive or significantly less negative than alphas generated by the least fluent portfolios.

To test for the possible impact of institutional ownership on stock returns, double sorted portfolios are also created based on one-year lagged, expanded, fluency score and both proxies for institutional ownership. To estimate levels of institutional ownership, two different variables are used, namely the percentage of total common stock in possession of institutions (IOTOT) and the Herfindahl-Hirschman Index of institutional ownership (OwnershipHI). The latter measurement is calculated as the sum of the squared market shares for each institution, per stock. Therefore, when fewer companies have higher stakes in a company, while still retaining high relative levels of institutional ownership, the higher the index.

Furthermore, I run cross-sectional regressions, inspired by Green & Jame (2003). These regressions are mainly to look at the influence of company name fluency and institutional ownership on firm value, but they also illustrate a relationship between fluency and institutional ownership on a cross-sectional level. To achieve this, I regress two different proxies for firm value, namely Tobin's Q and the Market-to-book ratio, on fluency, institutional ownership, multiple macro-economic variables and firm characteristics. The model looks as follows:

$$
\begin{aligned}
Firmvalue_{i,t} = {} & a_1\, Fluency_{it-1} + a_2\, Institutional\ Ownership_{it-1} + a_3 \log(Sales)_{it-1} \\
& + a_4\, Profitability_{it-1} + a_5 \log(Age)_{it-1} + a_6\, Salesgrowth_{it-1} + a_7\, Turnover_{it-1} \\
& + a_8 \left(\frac{R\&D}{Sales}\right)_{it-1} + a_9 \left(\frac{Adv}{Sales}\right)_{it-1} + a_{10} \log(Leverage) + a_{11}\, Payoutrate_{it-1} + \varepsilon_{it}
\end{aligned}
$$

In this model, firmvalue stands for the value of firm i at time t. This equation has multiple iterations, in either iteration, all independent variables are lagged one year. Variables followed by a_2 through a_{11} denote respectively; one of the two institutional ownership proxies (IOTOT or OwnershipHI, the logarithm of yearly sales (Sales), the profitability of a company (Profitability), the total years of existence of a company (Age), the yearly growth of sales (Salesgrowth), the yearly asset turnover (Turnover), the research and development- (R&D) and advertisement expenses (Adv) per unit of sales, the logarithm of the company leverage (log(Leverage) and the dividend payout ratio (Payoutrate).

The variable that requires a bit more in-depth coverage is the profitability measure. This variable originates from Compustat and is measured as the operating income before depreciation over the lagged total assets.

CHAPTER 5 Results

This section is twofold, and consists of both portfolio analysis and cross-sectional regressions. In the portfolio analysis, the goal is to estimate the effect of a stock's fluency score on the risk-adjusted returns. The goal of the cross-sectional regressions is to assess the effect of both fluency and institutional ownership on firm value.

5.1 Descriptive statistics fluency

This section shows the descriptive statistics of the returns, minus the risk-free rate, for each fluency score category. Additionally, they show the distributions for the fluency scores itself. These simple statistics give a good illustration of the averages, the deviations, the range and sample size per fluency category. This data serves as a very good starting point for research, as it is a good illustration of the data distribution and hints at possible relationships and their sign. The descriptive statistics also serve as a good check for the quality of the data.

Table 1. Descriptive statistics per fluency score before inclusion alphabet score

This table shows the descriptive statistics of stock returns for each fluency score category. This is <u>*before*</u> *expanding the fluency score measurement with the alphabet score. The results shown in the table are after adjustments to non-normally distributed variables and further cleansing of the data. For example, the statistics for score 0 are the descriptive statistics of the returns minus the risk-free rate for all stocks that have a fluency score of 0.*

Return		μ	σ	Min.	Max.	N
	Score	2.20	.877	0	4	554,416
	0	.006	.101	-.241	.272	15,977
	1	.006	.107	-.243	.272	95,696
Fluency	2	.005	.111	-.243	.272	213,264
	3	.005	.115	-.243	.272	196,421
	4	.007	.112	-.242	.272	20,770
	0-4	.006	.111	-.243	.272	542,128

Table 1. shows indication of a positive relationship between fluency score and returns. Even though the returns are not monotonically increasing over fluency scores, the main categories of interest are companies with fluency scores of zero and four. I start by performing a simple two-sample t-test, in which I compare the average returns of the least fluent companies (fluency=0) to the most fluent ones (fluency=4).

Table 2. Descriptive statistics per fluency score after inclusion alphabet score

This table shows the descriptive statistics of stock returns for each fluency score category. This is after expanding the fluency score measurement with the alphabet score. The results shown in the table are after adjustments to non-normally distributed variables and further cleansing of the data. For example, the statistics for score 0 are the descriptive statistics of the returns minus the risk-free rate for all stocks that have a fluency score of 0.

Return		μ	σ	Min.	Max.	N
	Score	2.37	.968	0	5	554,416
	0	.006	.103	-.241	.272	14,478
	1	.006	.107	-.243	.272	81,291
Fluency	2	.005	.110	-.243	.272	193,144
	3	.005	.114	-.243	.272	197,518
	4	.006	.117	-.243	.272	50,882
	5	.008	.110	-2.42	.272	4,815
	0-5	.006	.111	-.243	.272	542,128

Table 2. shows the exact same pattern as table 1. There are very slight differences however. Due to the creation of an extra fluency category, which contains all companies previously in category four with a name starting with letters a, b or c. The expansion of the fluency score this way increases the difference between the lowest and the highest fluency category ever so slightly. The only side-note here is that the sample size of the latter category is relatively small. There has also been a slight shift between categories, categories zero through three are rather similar whereas category four has gotten a drastic increase.

5.2 Portfolio analysis

I start by performing a simple two-sample t-test, in which I compare the average stock returns on the least fluent companies, before the addition of the alphabet score, to the most fluent ones. A one-sided two-sample t-test, on the difference in returns between fluency category four and zero leads to a t-statistic of 1.83. This implies that the most fluent companies generate, on average, higher returns than the least fluent companies.

Performing a one-sided t-test on the difference in stock returns between the most and least fluent companies, after the addition of the alphabet criterion, leads to a slightly more significant t-statistic of 2.41. This again implies that there is a significant positive difference in stock returns between the most- and the least-fluent companies. These results confirm what I hypothesized in sub-hypothesis 1.1., namely that there is a positive relationship between company name fluency and stock returns. In order to completely confirm the general hypothesis, the next step is to look at the risk-adjusted returns per fluency score.

Table 3. Portfolio analysis based on size and fluency score

*This table contains the equally-weighted risk-adjusted returns and t-statistics for each category based on expanded company-name fluency scores (ranging between 0 and 6) and deciles of size. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively.*

		SIZE									
		1	2	3	4	5	6	7	8	9	10
Fluency	0	0.008	0.012**	0.001	-0.003	0.004	0.007	0.006	0.005	0.005*	0.008**
		(1.52)	(2.54)	(0.26)	(-0.80)	(0.92)	(1.61)	(1.62)	(1.39)	(1.79)	(2.13)
	1	0.006	0.006*	0.004	0.006	0.007**	0.009***	0.006*	0.006**	0.007***	0.011***
		(1.49)	(1.68)	(0.97)	(1.62)	(2.02)	(2.73)	(1.75)	(2.08)	(2.60)	(3.04)
	2	0.005	0.002	0.005	0.004	0.005	0.007**	0.006**	0.007**	0.007***	0.010***
		(1.27)	(0.44)	(1.32)	(1.09)	(1.43)	(2.13)	(2.07)	(2.17)	(2.66)	(2.95)
	3	0.005	0.003	0.003	0.004	0.006	0.005	0.005	0.007**	0.007**	0.10***
		(1.22)	(0.69)	(0.91)	(1.14)	(1.58)	(1.45)	(1.45)	(2.29)	(2.30)	(2.96)
	4	0.007	0.002	0.003	0.004	0.004	0.006	0.006*	0.007**	0.006**	0.13***
		(1.62)	(0.53)	(0.67)	(1.01)	(1.01)	(1.45)	(1.72)	(2.03)	(1.99)	(3.19)
	5	0.007	0.011	0.007	0.004	0.018***	0.004	0.012*	0.010	0.010*	0.010**
		(0.53)	(1.54)	(0.91)	(0.47)	(3.30)	(0.81)	(1.66)	(1.32)	(1.83)	(2.54)

Judging solely from Table 3., there appears to be no clear-cut relationship between the company-name expanded fluency score and the risk-adjusted returns when controlling for size. Similarly, looking at the zero-investment portfolios, there is no definitive sign of significant risk-adjusted returns. For smaller companies (size categories one and two), the relationship appears to be slightly negative, however insignificant. Other categories, such as the largest companies, show an inverse relationship. For other categories, the relationship could be either positive or negative. Therefore, there appears to be no clear relationship between fluency score and risk-adjusted returns when correcting for size. However, when controlling for fluency, it is apparent that stocks of larger companies earn higher risk-adjusted returns.

Table 4. Portfolio analysis based on the Herfindahl-Hirschman index and fluency scores

*This table contains the value-weighted risk-adjusted returns and t-statistics for each category based on expanded company-name fluency scores (ranging between 0 and 6) and deciles of institutional ownership. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively. IOHI stands for the level of institutional ownership, using the Herfindahl-Hirschman index as a proxy.*

		IOHI									
		1	2	3	4	5	6	7	8	9	10
	0	0.009***	0.008**	0.006	0.009**	0.004	0.006*	0.005	0.005	0.004	0.000
		(3.05)	(2.17)	(1.59)	(2.23)	(1.23)	(1.67)	(1.30)	(1.06)	(0.85)	(-0.08)
	1	0.011***	0.011***	0.009***	0.007**	0.005	0.008**	0.005	0.004	0.006*	0.000
		(3.96)	(3.63)	(3.20)	(2.24)	(1.60)	(2.22)	(1.59)	(1.01)	(1.67)	(-0.03)
Fluency	2	0.010***	0.010***	0.009***	0.009***	0.008**	0.007**	0.005	0.003	0.002	-0.001
		(3.47)	(3.59)	(3.23)	(2.90)	(2.37)	(2.12)	(1.48)	(0.96)	(0.63)	(-0.38)
	3	0.010***	0.009***	0.009***	0.009***	0.006*	0.007**	0.004	0.002	0.002	-0.002
		(3.69)	(2.95)	(2.78)	(2.69)	(1.93)	(2.10)	(1.10)	(0.48)	(0.61)	(-0.46)
	4	0.008***	0.0121***	0.008**	0.011***	0.005	0.005	0.006	0.001	0.003	-0.002
		(2.64)	(3.40)	(2.08)	(3.30)	(1.30)	(1.31)	(1.58)	(0.31)	(0.73)	(-0.46)
	5	0.011***	0.007	0.000	0.012*	0.008	0.014**	0.005	0.016*	0.004	-0.009
		(2.79)	(1.22)	(-0.01)	(1.83)	(1.41)	(2.21)	(0.88)	(1.71)	(0.40)	(-0.86)

Controlling for a different variable in table 5., namely institutional ownership, there again appears to be no relationship between fluency scores and levels of risk-adjusted returns. There however does appear to be a negative relationship between the level of institutional ownership, using the Herfindahl-Hirschman index as a proxy, and risk-adjusted returns. This implies that companies with high levels of stock held by a very select few institutional investors, tend to generate lower risk-adjusted returns.

Table 5. Portfolio analysis based on total level of institutional ownership and fluency scores

*This table contains the value-weighted risk-adjusted returns and t-statistics for each category based on expanded company-name fluency scores (ranging between 0 and 6) and deciles of institutional ownership. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively. IOTOT stands for the level of institutional ownership, using the total percentage of shares held by institutions as a proxy.*

		IOTOT									
		1	2	3	4	5	6	7	8	9	10
	0	-0.003	0.009**	0.006**	0.005	0.006	0.005	0.009***	0.014***	0.002	0.005
		(-0.51)	(2.47)	(2.31)	(1.37)	(1.51)	(1.32)	(2.88)	(3.04)	(0.59)	(1.02)
	1	0.000	0.003	0.005	0.006*	0.007**	0.007**	0.009***	0.009***	0.010***	0.011***
		(-0.06)	(0.96)	(1.59)	(1.82)	(2.16)	(2.39)	(2.97)	(3.11)	(3.09)	(3.26)
Fluency	2	-0.001	0.003	0.004	0.007**	0.006**	0.008***	0.008***	0.008***	0.010***	0.010***
		(-0.41)	(0.96)	(1.27)	(2.45)	(2.08)	(2.60)	(2.63)	(2.77)	(3.20)	(3.03)
	3	-0.003	0.001	0.004	0.006*	1.71	0.005	0.009***	0.009***	0.010***	0.009***
		(-0.88)	(0.32)	(1.21)	(1.86)	(1.63)	(1.58)	(2.73)	(2.84)	(3.15)	(2.77)
	4	-0.001	0.002	0.005	0.007*	0.006*	0.006	0.008**	0.008**	0.009**	0.011***
		(-0.31)	(0.43)	(1.27)	(1.76)	(1.75)	(1.58)	(2.10)	(2.25)	(2.40)	(2.90)
	5	0.006	0.009	-0.009	0.019*	0.011*	0.004	0.005	0.013***	0.004	0.018
		(0.72)	(1.48)	(-0.86)	(1.84)	(1.80)	(0.61)	(1.00)	(2.93)	(0.79)	(1.62)

Similar to table 4., table 5. controls for levels of institutional ownership. The difference however is the use of a different proxy, namely the total percentage of shares held by institutional owners. Using this different proxy, the relationship between fluency scores and risk-adjusted returns does not seem to change and is still inconclusive. However, the opposite is true for institutional ownership. Rather than the negative relationship between institutional ownership derived from the other proxy, the use of the total percentage of shares held by institutional owners leads to a positive relationship between institutional ownership and risk-adjusted returns.

The results for table 5. are to be expected, as an institutional investor should possess more information on stocks and spend more time analyzing and buying them. Please note, that these returns do not account for transaction costs, which tend to be higher for some institutional investors categories, as they make more transactions and therefore have a higher transaction costs.

The results for table 4. are not as clear, however. These imply that higher levels of institutional ownership are not beneficial to stock returns. The difference between institutional ownership and the Herfindahl-Hirschman index lies in the construction of the Herfindahl-Hirschman index. Even when relative levels of institutional ownership are high, the Herfindahl-Hirschman index can take on relatively low values due to a large amount of institutional owners owning stock in a company.

Table 6. Portfolio analysis based on company size and names starting with letters A, B or C

*This table contains the equally-weighted risk-adjusted returns and t-statistics for each category based company names starting with either the letters a, b or c and deciles of size. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively.*

		SIZE									
		1	2	3	4	5	6	7	8	9	10
	ABC	0.006	0.003	0.003	0.003	0.005	0.006	0.006*	0.007**	0.006*	0.012***
		(1.59)	(0.68)	(0.87)	(0.77)	(1.40)	(1.72)	(1.74)	(2.30)	(1.86)	(3.55)
	A	0.004	-0.001	0.001	0.003	0.005	0.007	0.002	0.004	0.010**	0.010**
		(0.68)	(-0.21)	(0.08)	(0.53)	(0.80)	(1.03)	(0.31)	(0.70)	(2.38)	(2.47)
Alphabet	B	0.006	0.001	0.007	0.004	0.0147***	0.009**	0.008*	0.008**	0.003	0.012***
		(0.79)	(0.11)	(1.30)	(0.87)	(3.13)	(2.01)	(1.85)	(2.26)	(0.86)	(3.29)
	C	0.007*	0.003	0.005	0.003	0.004	0.006	0.006*	0.008**	0.005*	0.012***
		(1.85)	(0.80)	(1.18)	(0.76)	(1.04)	(1.65)	(1.80)	(2.33)	(1.75)	(3.54)

To look at the effect of the fluency score augmentation in isolation, table 6. Shows the returns for stocks starting with either letters A, B or C, both individually and in conjunction, while controlling for firm-size. The risk-adjusted returns for companies starting with letters early in the alphabet appear to be slightly positive, but not very statistically significant. For the largest ten percent of companies, the effect however seems fairly pronounced at around 1.2% monthly excess returns.

To summarize the main results of the portfolio analyses, regarding fluency scores, for none of the portfolio specifications do I find a positive relationship between company name fluency and excess returns. This goes against what I hypothesized in sub-hypothesis 1.2, namely that the excess return generated by more fluently named stocks cannot be classified as a compensation for increased risk exposure. Therefore, since hypothesis 1.2 does not hold true, neither does hypothesis 1. To conclude, more fluently named company stock does not seem to outperform less fluently named company stock on the North-American stock market (between 1983 and 2008).

Additionally, I find that higher levels of total institutional ownership lead to, on average, significantly higher excess stock returns. Contrarily, judging solely from the used institutional ownership index, I find a negative effect of institutional ownership on fluency. Due to the previously explained difference between the two proxies for institutional ownership, I still confirm what I hypothesized in hypothesis 3., namely that companies with higher levels of institutional ownership generate higher stock returns, disregarding transaction costs. As explained before, this however does not hold true, or to a lesser extent, whenever very few institutions own large clusters of the particular stock.

5.3 Cross-sectional analysis

To give a first impression of the data that is used in the cross-sectional regressions, table 7. Shows descriptive statistics of all variables, and their specification, that are used in the cross-section.

Table 7. Descriptive statistics of cross-sectional variables

This table contains the descriptive statistics of each variable after adjustments. Possible adjustments to variables are the following: (1) the deletion of negative (false) values for variables which are strictly positive by definition. (2) taking the natural logarithm of a variable. (3) winsorization of the variable at the 1^{st}, 5^{th} or 10^{th} percentile.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.	Jarque-Bera
Tobin's Q	383,488	1.636	0.932	0.763	4.389	2.5×10^5
Log(M2B)	392,639	0.592	0.752	-0.865	2.164	3.3×10^3
Fluencyscore	789,448	2.192	0.875	0.000	4.000	1.6×10^4
Lengthscore	789,448	1.067	0.705	0.000	2.000	3.3×10^4
Dictionaryscore	789,448	0.312	0.464	0.000	1.000	1.5×10^5
Englishness-score	789,448	0.811	0.391	0.000	1.000	3.4×10^5
Ownershiptot	742,467	0.756	20.492	1.5×10^{-8}	10.8×10^2	2.4×10^{11}
OwnershipHI × Fluencyscore	625,580	0.437	0.565	0.000	4.000	2.0×10^6
Ownershiptot × Fluencyscore	622,374	1.907	51.373	0.000	32.4×10^2	1.9×10^{11}
Log(sales)	373,790	12.887	1.992	9.310	16.398	1.2×10^4
Profitability	330,753	0.387	0.375	-0.196	1.357	6.7×10^4
Log(age)	946,882	2.599	0.906	0.000	4.419	3.9×10^4
Salesgrowth	363,713	0.416	0.591	-0.337	2.083	1.5×10^5
Turnover	366,360	1.096	0.690	0.094	2.652	2.2×10^4
RD/sales	190,490	0.061	0.082	0.000	-0.312	1.6×10^5
AD/sales	119,596	0.027	0.029	0.001	0.110	7.5×10^4
Log(leverage)	125,317	1.939	1.120	0.552	5.103	3.7×10^4
Payoutrate	371,462	0.219	0.293	0.000	0.986	1.1×10^5

Table 7. shows the Jarque-Bera statistic, which estimates the normality of each variable used in the cross-sectional regressions. This statistic clearly shows that none of these variables are statistically normally distributed, however, making more adjustments to these variables than logarithmic transformations, wherever possible, and winsorization at the 1^{st}, 5^{th} or 10^{th} percentile

is not justifiable. The final iterations of each variable are made based on box-plots, graphed for each possible iteration of the variable (non-log winsorized, log winsorized etcetera), and stopping whenever each variable looks approximately normally-distributed from a box-plot, with as little adjustments as possible.

The following step is to actually run these cross-sectional regressions. table 8. Shows cross-sectional regressions on both variables for firm value, for both proxies of institutional ownership. Additionally, regressions five and six clearly illustrate the relationship between fluency scores, institutional ownership and the used control variables.

Table 8. Cross-sectional regressions

*Regression equation (1) and (2) regress Tobin's Q on the listed independent variables. Regressions equations (3) and (4) regress the logarithm of the Market-to-book ratio on the same variables. Regressions (2) and (4) include the percentage of common stock owned by institutional investors, proxying for levels of institutional ownership per stock, whereas regressions (1) and (3) use the Herfindahl-Hirschman index to proxy for levels of institutional ownership. In regressions (5) and (6), the dependent variable is the fluency score of each stock. Dep. implies that this variable is the dependent variable in this regression and – implies that these variables are left out of the regression completely. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively*

	(1)	**(2)**	(3)	**(4)**	**(5)**	**(6)**
Fluency	0.097***	0.037***	0.104***	0.036***	Dep.	Dep.
	(11.63)	**(3.05)**	**(15.03)**	**(3.59)**		
OwnershipHI	0.311***	-	0.341***	-	-	-4.611***
	(4.34)		**(5.75)**			**(-83.19)**
Ownershiptot	-	-0.027	-	-0.019	-3.267***	-
		(-0.51)		**(-0.43)**	**(-138.83)**	
Ownership ×	-0.376***	0.007	-0.352***	0.032	1.579***	2.33
Fluency	**(-11.93)**	**(0.31)**	**(-13.45)**	**(1.71)**	**(194.42)**	**(102.26)**
Log(sales)	0.074***	0.106***	0.064***	0.085***	0.009***	0.013***
	(16.43)	**(24.85)**	**(16.92)**	**(24.00)**	**(3.12)**	**(3.11)**
Profitability	0.691***	0.664***	0.413***	0.390***	0.087***	0.053
	(31.36)	**(24.85)**	**(22.58)**	**(21.24)**	**(5.92)**	**(2.55)**
Log(age)	0.023***	0.015*	0.059***	0.055***	-0.026***	0.050***
	(2.72)	**(1.74)**	**(8.26)**	**(7.55)**	**(-4.63)**	**(6.21)**
Salesgrowth	0.340***	0.393***	0.226***	0.022***	-0.027***	-0.059***
	(31.54)	**(30.91)**	**(21.60)**	**(21.01)**	**(-3.26)**	**(-4.95)**
Turnover	0.184***	0.176***	0.087***	0.079***	-0.069***	-0.064***
	(14.61)	**(30.91)**	**(8.27)**	**(7.45)**	**(-8.40)**	**(-5.52)**
R&D/Sales	5.500***	5.610***	3.468***	3.531***	0.138*	0.418***
	(51.10)	**(51.57)**	**(38.86)**	**(39.13)**	**(1.83)**	**(3.93)**

AD/Sales	4.922***	5.118***	3.462***	3.657**	-4.70***	-0.805***
	(21.67)	**(22.37)**	**(18.38)**	**(19.27)**	**(-3.15)**	**(-3.80)**
Log(leverage)	0.112***	0.115***	-0.230***	-0.027***	0.028***	0.035***
	(19.78)	**(20.08)**	**(6.24)**	**(-5.74)**	**(7.23)**	**(6.31)**
Payoutrate	0.233***	0.221***	0.275***	0.265	-0.151***	-0.438***
	(8.65)	**(8.19)**	**(12.32)**	**(11.83)**	**(-8.69)**	**(-17.81)**
Constant	-0.641	-0.980***	-0.983***	-1.193***	2.050***	1.763***
	(-9.53)	**(-15.97)**	**(-17.62)**	**(-23.42)**	**(55.60)**	**(28.87)**
Tobin’s Q	Dep.	Dep.	-	-	0.001	-0.020***
					(0.12)	**(-5.10)**
Log(Market/Book)	-	-	Dep.	Dep.	0.020	0.148***
					(1.88)	**(9.60)**
Obs.	16,693	16,670	16,693	16,670	16,670	16,693
R^2	0.325	0.315	0.250	0.238	0.708	0.414
Adj. R^2	0.324	0.314	0.249	0.237	0.708	0.413
F-stat	668.47	637.61	463.25	433.17	3101.87	904.79

The first regression in table 8., using the Herfindahl-Hirschman index as a proxy for levels of institutional ownership and Tobin’s Q as a proxy for firm value, shows a significantly positive effect of fluency and institutional ownership on firm value. The interaction-effect between the two variables, however, is negative, and also more negative than the positive coefficient for institutional ownership. Therefore, For all fluency scores other than zero, the effect of an increase in the institutional ownership index is net-negative, and decreases with each increase in fluency score. Similarly, this interaction effect diminishes the positive effect of fluency on firm value. However, an institutional ownership index of larger than 0.25 is quite rare and therefore the effect of fluency on firm value will on average be positive.

Using the total percentage of stocks owned by institutions as a proxy for institutional ownership in regression (2) of table 8., all else equal, the significant effect of fluency on firm value remains. The isolated effect of institutional ownership on firm value, however, is negative. This changes to net-positive for higher fluency scores but remains insignificant for each fluency score. Due to positivity of the interaction term, the effect of fluency score on firm value remains positively significant for each level of institutional ownership.

Regression (3) of table 8., again using the institutional ownership index as a proxy for levels of institutional ownership, shows the exact same effect for both fluency and institutional ownership on firm value as in regression (1). These results make sense, as the only difference between these two models is the proxy for firm value, therefore, these results seem robust across both proxies for firm value.

Regression (4) of table 8. Stands to regression (2), as regression (3) stands to regression (1). The only difference here is the use of the Market-to-Book ratio as a proxy for firm value rather than Tobin's Q. There is one difference between the results of regression (4) and (2), which is that regression (4) shows a positively significant effect of institutional ownership on firm value for fluency scores of two and higher, which is the majority of the sample. The effect of fluency on firm value remains positive for each level of institutional ownership, just like in regression (2).

Lastly, regressions (5) and (6) regress all variables on fluency score, rather than regressing fluency score, institutional ownership and all control variables on proxies for firm value. These regressions are mostly for illustrative purposes and robustness, as this shows the relationships between institutional ownership, fluency and the variables of control used in the regressions.

To summarize the results of all regressions included in table 8., the effect of fluency of firm value is, on average, significantly positive for all model specifications. A higher institutional ownership index (Herfindahl-Hirschman) has a net-negative effect on both proxies for firm value for the majority of the sample data. Contrarily, the effect of a higher percentage of total institutional ownership on firm value is net positive, however, the results are not very robust across both proxies for firm value and for each fluency score. These results therefore confirm what I hypothesized in hypothesis 2., which states that higher levels of company name fluency lead to higher levels of firm value.

In order to fully clarify the effect of institutional ownership on firm value, both used proxies for institutional ownership require more thorough explanation.

The Herfindahl-Hirschman index captures the extent to which common-stock of a specific company is held by large blocks of investors. In practice, this variable takes on high values, when one, or very few, investors hold the majority of shares. Therefore, it could be that the total percentage of institutional ownership is high, while the Herfindahl-Hirschman index of institutional ownership is relatively low. This occurs when many institutions hold fewer shares in a company, rather than one institution holding all of the shares, while still retaining high levels of institutional ownership.

The other variable used to proxy for levels of institutional ownership is the total percentage of stock held by institutional investors. This variable specification is fairly intuitive and requires very few explanation. In short, if 75% of stock is held by institutional investors, this variable takes on the value of 0.75, regardless of the total amount of institutional investors holding these stocks.

Therefore, we can draw from this, that higher levels of institutional ownership are beneficial to firm value. However, when these high levels of institutional ownership are due to a small number of institutional investors holding the majority of shares, the effect is negative. All in all, the best case scenario in terms of maximizing firm value, would be to have as many institutional investors hold as few shares as possible, while still maintaining high levels of total institutional ownership. This somewhat confirms what I hypothesized in hypothesis 4. Namely, that companies with higher levels of institutional ownership have higher firm values. However, the real answer is more nuanced, as explained.

CHAPTER 6 Conclusion

I find empirical evidence for the existence of a fluency effect on the North-American stock market between 1983-2008. However, like most other investment strategies, the on average higher returns from this strategy seem to essentially be a compensation for higher levels of risk exposure.

While performing portfolio analysis, I find that higher levels of institutional ownership lead to, on average, higher risk-adjusted returns. However, the relationship between institutional ownership and stock returns is more nuanced. When institutional ownership is highly concentrated, meaning that few companies hold lots of stock, risk-adjusted returns are negative. Therefore, when trying to account for institutional ownership in investment decisions, ownership concentration should be taken into consideration

The results are similar in the cross-section when looking at the impact of institutional ownership on the two proxies for firm value, namely Tobin's Q and the Market-to-book ratio. Higher levels of institutional ownership lead to higher levels of firm value, whereas large institutional ownership concentration, proxied by the Herfindahl-Hirschman index, leads to lower firm values.

Empirical evidence shows that company name fluency has a positive effect on both proxies for firm value in the cross-section, while controlling for institutional ownership and multiple other, mostly macro-economic, variables. This is explainable by the argument of Head, Smith & Wilson (2003), who state that fluency is related to managerial ability, due to capable managers being more likely to opt into a more fluent company name.

Looking at the isolated effect of the starting letter for each company name on risk-adjusted stock returns, the results appear to be marginally statistically significant, but highly economically significant. The effect however is very pronounced for the largest decile of stocks. This leads me to believe that the alphabet score serves as an appropriate additional parameter for fluency. However, due to the lack of robustness-checks and the small level of statistical significance, it is risky to draw further conclusions from these results alone. It does however look very interesting for further research to look into the impact of the alphabet-effect on stock returns, mainly for large companies.

Another recommendation for further research would be to further look into the effect of institutional ownership on both stock returns and firm value. The empirically observed relationship between the total amount of institutional ownership in percentages, and the institutional ownership concentration, is particularly interesting. Mapping the relationship between these two variables, either graphically or mathematically, and their impact on stock returns on firm value might lead to interesting results.

A limitation to my research is the lack of data or research on the impact of name changes on stock returns. Performing a case-study on the impact of a name-change on risk-adjusted stock returns might show a relationship between the impact of company name fluency and stock returns. Safe to say, there are plenty of opportunities for further research involving both company name fluency, institutional ownership and the alphabet-effect

REFERENCES

Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences*, *103*(24), 9369-9372.

Alter, A. L., & Oppenheimer, D. M. (2008). Easy on the mind, easy on the wallet: The roles of familiarity and processing fluency in valuation judgments. *Psychonomic Bulletin & Review*, *15*(5), 985-990.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219-235.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31-56.

Asquith, P., Pathak, P. A., & Ritter, J. R. (2005). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics*, *78*(2), 243-276.

Blume, M. E., & Keim, D. B. (2014). The changing nature of institutional stock investing. *Critical Finance Review*, *6*, 1-41.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Chan, C. S. R., Park, H. D., & Patel, P. (2018). The effect of company name fluency on venture investment decisions and IPO underpricing. *Venture Capital*, *20*(1), 1-26.

Chung, K. H., & Pruitt, S. W. (1994). A simple approximation of Tobin's q. *Financial Management*, 70-74.

Coval, J. D., & Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance*, *54*(6), 2045-2073.

Dreman, D. (2004). The influence of affect on investor decision-making. *The Journal of Behavioral Finance*, *5*(2), 70-74.

Edelen, R. M., Ince, O. S., & Kadlec, G. B. (2016). Institutional investors and stock return anomalies. *Journal of Financial Economics*, *119*(3), 472-488.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.

Fama, E. F., & French, K. R. (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies*, *29*(1), 69-103.

Ferson, W. E., Sarkissian, S., & Simin, T. (1999). The alpha factor asset pricing model: A parable. *Journal of Financial Markets*, *2*(1), 49-68.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451-482.

Gompers, P. A., & Metrick, A. (2001). Institutional investors and equity prices. *The Quarterly Journal of Economics*, *116*(1), 229-259.

Green, T. C., & Jame, R. (2013). Company name fluency, investor recognition, and firm value. *Journal of Financial Economics*, *109*(3), 813-834.

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity. *The Review of Financial Studies*, *17*(2), 439-461.

Head, A., Smith, G., & Wilson, J. (2009). Would a stock by any other ticker smell as sweet?. *The Quarterly Review of Economics and Finance*, *49*(2), 551-561.

Holcomb, T. R., Holmes Jr, R. M., & Connelly, B. L. (2009). Making the most of what you have: Managerial ability as a source of resource value creation. *Strategic Management Journal*, *30*(5), 457-485.

Huang, P. H. (2005). Moody investing and the Supreme Court: Rethinking the materiality of information and the reasonableness of investors. *Supreme Court Economic Review*, *13*, 99.

Hung, W. (2014). Institutional trading and attention bias. *Journal of International Financial Markets, Institutions and Money*, *29*, 71-91.

Itzkowitz, J., Itzkowitz, J., & Rothbort, S. (2015). ABCs of trading: Behavioral biases affect stock turnover and value. *Review of Finance*, *20*(2), 663-692.

Li, X., Mahani, R., & Sandhya, V. (2011). Does Investor Attention Affect Stock Prices?.

Pensa, P. (2006). Nomen est Omen: How Company Names Influence Short-and Long-Run Stock Market Performance.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Van den Assem, M. J., Montone, M., & Zwinkels, R. C. (2018). Company Name Fluency and Stock Returns.

Xing, X., Anderson, R. I., & Hu, Y. (2016). What's a name worth? The impact of a likeable stock ticker symbol on firm value. *Journal of Financial Markets*, *31*, 63-80.

Yan, X., & Zhang, Z. (2007). Institutional investors and equity returns: Are short-term institutions better informed?. *The Review of Financial Studies*, *22*(2), 893-924.

APPENDIX

This appendix contains all 6x10 tables based on the fluency score expanded with the alphabet score and deciles for institutional ownership proxies or size. Most of these tables are not included in the main text due to them serving the same purpose, with nearly equal results, as the tables included in the main text. These tables serve as a robustness check or as extra information for the in-depth reader, and are therefore included in the appendix. Additionally, the last table in the appendix is a correlation matrix of all variables used in the cross-sectional regressions.

Table 9. Portfolio analysis based on the Herfindahl-Hirschman index and fluency scores (equally-weighted)

*This table contains the equally-weighted risk-adjusted returns and t-statistics for each category based on expanded company-name fluency scores (ranging between 0 and 6) and deciles of institutional ownership. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively. IOHI stands for the level of institutional ownership, using the Herfindahl-Hirschman index as a proxy.*

		IOHI									
		1	2	3	4	5	6	7	8	9	10
	0	0.010***	0.008**	0.006	0.009**	0.004	0.006	0.005	0.004	0.003	-0.001
		(3.18)	(2.12)	(1.52)	(2.17)	(1.18)	(1.55)	(1.15)	(0.96)	(0.75)	(-0.22)
	1	0.011***	0.011***	0.009***	0.007**	0.004	0.007**	0.005	0.002	0.005	-0.001
		(3.93)	(3.56)	(3.08)	(2.08)	(1.35)	(2.03)	(1.38)	(0.67)	(1.37)	(-0.29)
Fluency	2	0.010***	0.010***	0.009***	0.008***	0.007**	0.006*	0.004	0.002	0.001	-0.003
		(3.41)	(3.52)	(3.06)	(2.73)	(2.15)	(1.85)	(1.21)	(0.64)	(0.28)	(-0.74)
	3	0.010***	0.009***	0.008***	0.009**	0.006*	0.006*	0.003	0.001	0.001	-0.003
		(3.60)	(2.84)	(2.61)	(2.50)	(1.70)	(1.83)	(0.84)	(0.14)	(0.27)	(-0.81)
	4	0.008**	0.12	0.007*	0.010***	0.004	0.004	0.005	0.000	0.002	-0.004
		(2.51)	(3.22)	(1.89)	(3.14)	(1.07)	(1.06)	(1.40)	(-0.03)	(0.49)	(-0.93)
	5	0.011***	0.007	-0.001	0.011*	0.008	0.014**	0.005	0.016*	0.004	-0.009
		(2.78)	(1.21)	(-0.12)	(1.77)	(1.32)	(2.16)	(0.83)	(1.71)	(0.38)	(-0.86)

Table 10. Portfolio analysis based on total level of institutional ownership and fluency scores (equally-weighted)

*This table contains the equally-weighted risk-adjusted returns and t-statistics for each category based on expanded company-name fluency scores (ranging between 0 and 6) and deciles of institutional ownership. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively. IOTOT stands for the level of institutional ownership, using the total percentage of shares held by institutions as a proxy.*

		IOTOT									
		1	2	3	4	5	6	7	8	9	10
	0	-0.003	0.009**	0.006**	0.005	0.005	0.004	0.008**	0.014***	0.002	0.005
		(-0.66)	(2.25)	(2.09)	(1.30)	(1.22)	(1.23)	(2.53)	(3.06)	(0.48)	(0.96)
	1	-0.001	0.002	0.004	0.005	0.006*	0.007**	0.009***	0.009***	0.010***	0.011***
		(-0.37)	(0.63)	(1.26)	(1.53)	(1.89)	(2.24)	(2.77)	(2.93)	(2.96)	(3.21)
Fluency	2	-0.003	0.002	0.003	0.007**	0.005*	0.007**	0.007**	0.008**	0.010***	0.010***
		(-0.76)	(0.57)	(0.89)	(2.16)	(1.74)	(2.37)	(2.40)	(2.52)	(3.04)	(2.89)
	3	-0.005	0.000	0.003	0.005	0.005	0.004	0.008**	0.008***	0.009***	0.009***
		(-1.21)	(-0.05)	(0.87)	(1.58)	(1.42)	(1.29)	(2.54)	(2.63)	(2.95)	(2.62)
	4	-0.003	0.000	0.003	0.006	0.006	0.005	0.007*	0.008**	0.008**	0.010***
		(-0.68)	(0.01)	(0.90)	(1.50)	(1.54)	(1.25)	(1.85)	(2.06)	(2.22)	(2.69)
	5	0.006	0.008	-0.009	0.019*	0.011*	0.004	0.004	0.013***	0.004	0.018
		(0.66)	(1.35)	(-0.90)	(1.78)	(1.74)	(0.56)	(0.84)	(2.85)	(0.81)	(1.62)

Table 11. Portfolio analysis based on size and fluency score (value-weighted)

*This table contains the value-weighted risk-adjusted returns and t-statistics for each category based on expanded company-name fluency scores (ranging between 0 and 6) and deciles of size. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively.*

		SIZE									
		1	2	3	4	5	6	7	8	9	10
Fluency	0	0.009	0.013***	0.002	-0.003	0.004	0.008*	0.007*	0.005	0.005*	0.008**
		(1.63)	(2.79)	(0.42)	(-0.61)	(1.06)	(1.76)	(1.73)	(1.51)	(1.84)	(2.16)
	1	0.007*	0.008**	0.005	0.007*	0.008**	0.009***	0.006*	0.007**	0.007***	0.012***
		(1.80)	(2.03)	(1.33)	(1.94)	(2.30)	(3.01)	(1.95)	(2.24)	(2.72)	(3.08)
	2	0.006	0.003	0.006*	0.005	0.006*	0.008**	0.007**	0.007**	0.008***	0.010***
		(1.60)	(0.89)	(1.71)	(1.45)	(1.75)	(2.39)	(2.30)	(2.36)	(2.77)	(2.99)
	3	0.006	0.005	0.004	0.005	0.005	0.007*	0.006*	0.008**	0.007**	0.010***
		(1.54)	(1.42)	(1.12)	(1.29)	(1.48)	(1.85)	(1.66)	(2.49)	(2.43)	(3.02)
	4	0.008**	0.004	0.004	0.005	0.005	0.007*	0.007	0.008**	0.007**	0.013***
		(1.99)	(0.94)	(1.63)	(1.20)	(1.32)	(1.68)	(1.95)	(2.20)	(2.10)	(3.23)
	5	0.007	0.011	0.008	0.004	0.018***	0.005	0.012*	0.010	0.010*	0.010**
		(0.52)	(1.59)	(1.05)	(0.52)	(3.37)	(0.89)	(1.67)	(1.35)	(1.84)	(2.56)

Table 12. Portfolio analysis based on company size and names starting with letters A, B or C (value-weighted)

*This table contains the value-weighted risk-adjusted returns and t-statistics for each category based company names starting with either the letters a, b or c and deciles of size. The listed percentages are alphas from Carhart 4-factor model regressions. The student t-statistic is presented in parentheses, and *, **, *** indicate levels of significance of 10%, 5% and 1% respectively.*

		SIZE									
		1	2	3	4	5	6	7	8	9	10
	ABC	0.007*	0.004	0.005	0.004	0.006*	0.007**	0.007**	0.008**	0.006**	0.012***
		(1.91)	(1.13)	(1.29)	(1.13)	(1.69)	(1.98)	(1.99)	(2.41)	(1.99)	(3.59)
Alphabet	A	0.005	0.000	0.002	0.004	0.006	0.008	0.002	0.005	0.010**	0.010**
		(0.81)	(0.02)	(0.32)	(0.68)	(0.94)	(1.22)	(0.43)	(0.79)	(2.44)	(2.50)
	B	0.007	0.002	0.008	0.005	0.015***	0.009**	0.008*	0.008**	0.003	0.012***
		(0.92)	(0.31)	(1.51)	(1.04)	(3.25)	(2.15)	(1.96)	(2.36)	(0.93)	(3.31)
	C	0.009**	0.005	0.006	0.004	0.005	0.007*	0.007**	0.008**	0.006*	0.012***
		(2.16)	(1.22)	(1.55)	(1.11)	(1.33)	(1.90)	(2.05)	(2.55)	(1.89)	(3.58)

Table 13. Correlation matrix cross-sectional variables

This table shows the correlation between all variables that are used in the cross-sectional regressions. These results show how closely related each variable is to another.

Correlation Matrix	Tobin's Q	Log(M2B)	Fluency	OwnershipTOT	OwnershipHI	OwnershipHI × Fluency	OwnershipTOT × Fluency	Log(Sales)	Profitability	Log(Age)	Salesgrowth	Turnover	RD/Sales	AD/Sales	Log(Leverage)	Payoutrate
Tobin's Q	1.000															
Log(M2B)	0.848	1.000														
Fluency	0.097	0.093	1.000													
OwnershipTOT	0.136	0.169	0.036	1.000												
OwnershipHI	-0.023	-0.255	-0.049	-0.602	1.000											
OwnershipHI × Fluency	-0.191	-0.226	0.290	-0.517	0.846	1.000										
OwnershiTOT × Fluency	0.154	0.183	0.556	0.775	-0.470	-0.312	1.000									
Log(Sales)	0.206	0.290	0.003	0.480	-0.573	-0.509	0.366	1.000								
Profitability	0.188	0.215	0.019	0.160	-0.150	-0.114	0.128	0.403	1.000							
Log(Age)	0.001	0.115	-0.013	0.077	-0.124	-0.137	0.071	0.396	0.135	1.000						
Salesgrowth	0.241	0.1338	0.030	-0.022	0.038	0.060	-0.000	-0.116	0.028	-0.299	1.000					
Turnover	-0.089	-0.091	-0.090	0.0417	0.048	0.010	-0.005	0.112	0.109	0.014	-0.034	1.000				
RD/Sales	0.308	0.188	0.131	-0.008	-0.003	0.040	0.061	-0.245	-0.334	-0.196	0.148	-0.498	1.000			
AD/Sales	0.179	0.195	-0.033	0.023	-0.128	-0.132	0.000	0.160	0.154	0.121	-0.046	-0.097	-0.034	1.000		
Log(Leverage)	0.148	-0.059	0.083	-0.066	0.061	0.086	-0.014	-0.233	-0.251	-0.150	0.073	-0.019	0.287	-0.095	1.000	
Payoutrate	0.037	0.133	-0.108	0.081	-0.176	-0.172	0.010	0.330	0.209	0.358	-0.223	-0.011	-0.219	0.152	-0.157	1.000