The effect of fluency across industries and firm types

Nicolette van den Bosch (382997)

June 20, 2018

Master's Thesis

Erasmus School of Economics

Supervisor: Dr M. Montone

Abstract

This paper examines whether the fluency effect varies across different industries and different types of firms. The results show that effect of fluency is more pronounced in industrial manufacturing sectors than in other industries. Moreover, this paper finds that the fluency effect is the strongest for older firms and firms with a high market capitalization, a low book-to-market ratio and a high level of intangible assets. The results imply that investors may be able to exploit the fluency effect by concentrating their investments in stocks of firms with a fluent name that operate in industrial manufacturing sectors and have characteristics for which the effect is stronger.

Contents

1 Introduction

According to traditional finance theory, investors incorporate all public and private information to make investment decisions. When choosing between the thousands of available stocks, they use this information to construct the optimal portfolio (Fama, 1970). However, a recent line of academic papers challenges this rational framework. According to behavioral finance, investors do not always behave as rationally as the traditional finance suggests. The advocates of this field of finance argue that human psychology influences investment decisions.

A behavioral bias that recently receives the attention of academics is the fluency effect. This means that certain stocks are more likeable because they have a name that is easy to process by the human brain (Alter and Oppenheimer, 2008). Previous literature shows that the fluency bias affects the stock market. However, not much is known about the variance of this effect across different environments. This paper adds to the existing literature by studying the effect of fluency on stock returns. In particular, I focus on the differences of the fluency effect on stock returns across different industries and firms with different characteristics.

This study assigns fluency scores to company names based on the methodology developed by Green and Jame (2013). This methodology proposes that processing fluency depends on word length and the ease of pronunciation. The overall fluency score is the aggregate of three separate measures that capture word length and pronunciation, namely the length, Englishness and dictionary scores. Based on these three dimension, firms obtain a fluency score that ranges from 1 to 5.

Using data on US firms from 1988 to 2017, this study first tests whether fluency has an effect on excess returns on the full sample. I test this by using both a simple regression without any controls and a multiple panel regression that incorporates control variables. The estimates of these regression indicate that fluency positively affects excess returns. The results hold after controlling for firm characteristics that are, according to Hong and Kacperczyk (2009), determinants of return.

After this initial test, the sample is categorized into 48 subsamples. This categorization is based on the Fama and French (1997) 48-industry classfication. Regressions of excess returns on fluency are performed for each of the 48 subsamples. This process is repeated for a test that includes the same control variables that are used in the first test. The estimates of these tests indicate that the fluency effect varies across different industries. The effect is both economically and statistically significant for the tobacco, fabricated metal products, container, and hotel and restaurant industries. For all other industries, the coefficients of the regression are statistically not different from zero.

To test whether the fluency effect varies across firms with different characteristics, the firms are sorted into quintiles based on these characteristics. For each quintile of each characteristic, a regression of excess returns on fluency is performed. Furthermore, I perform a regression including interaction terms between fluency and the characteristic of interest is performed. This way, the relation between fluency and the characteristics can be analyzed using interaction effects.

The results indicate that the fluency effect differs across firms with different types of characteristics. Processing fluency is increasing in size. The cause of this result may be the more intense analyst coverage on larger firms, which reduces the possibility of a mispricing such as the fluency effect.

Moreover, the bias is more pronounced in firms with a low book-to-market ratio. A low book-to-market ratio is considered to be an indicator of mispricing (Lakonishok et al., 1994). Therefore, the fluency effect is stronger in firms with a low book-to-market ratio.

Moreover, older firms are more susceptible to the fluency effect. This may be explained by the survivorship bias. However, this result remains unexpected.

Finally, the fluency effect is more pronounced in firms with a high amount of intangible assets. Because the information provision of firms with a high level of intangible asset is less efficient, the reduced information availability makes a firm more sensitive to mispricing. Therefore, the fluency effect biases the stock returns of firms with a lot of intangible assets.

The fluency effect does not differ across firms with different levels of asset liquidity and book leverage.

1.1 Literature review

Other literature that also applies psychology to explain the stock market, focuses on the human tendency to make use of heuristics. This line of literature argues that investors do not take all available information into account. Instead, they are influenced by cognitive limitations that may cause their decisions to be suboptimal. These cognitive limitations often result in the use of heuristics. Because of this, heuristics play a significant role on financial markets.

Tversky and Kahneman (1973) propose people use a limited amount of heuristics when they assess difficult situations. This contradicts the view of traditional finance that all investors are rational when they make their portfolio choices. Instead of the fully informed decisions investors make in the efficient market hypothesis, Tversky and Kahneman (1973) suggest that people use mental shortcuts when making choices. This results in decisions that are made based on, for example, the representativeness and the availability of information. This leads to cognitive biases such as the availability bias.

Another bias that may be the result of mental shortcuts, is the familiarity bias. This bias indicates that investors simply prefer that what is known to them. It reflects people's tendency to be optimistic about what they feel affinity with. Based on the degree of affinity, this may cause investors to focus on familiar options and ignore unfamiliar options that may be a better choice. (Huberman, 2001). This degree of familiarity depends on a number of factors, such as the geographical and professional proximity of the option to the agent (Massa and Simonov, 2006).

Reber et al. (2004) indicate that an additional factor, that can influence investors' evaluation and decision making process, is aesthetic pleasure. They show that objects that can be more fluently processed receive a more positive response. Multiple studies support the finding that fluency affects evaluation and decision making. Song and Schwarz (2009) find that amusement park rides with more

difficult names are deemed to be riskier than those with easier names. The same authors find that instructions written in less fluent language are perceived to be more difficult. Therefore, people are less willing to participate in tasks that involve instructions that are difficult to interpret (Song and Schwarz, 2008).

Processing fluency also affects product and company performance. Bao et al. (2008) show that products with easier names have higher brand recognition. Another study indicates that name changes can influence the decisions of investors (Cooper et al., 2005). Begg et al. (1992) show that investors are more likely to think that the content of a company's statement is true when is written in language that is easy to understand.

Recent literature shows that fluency also influences the stock market. Alter and Oppenheimer (2008) argue that processing fluency can cause an asset to become more familiar to a person. This results in a bias that causes fluent stocks to be more valued than less fluent stocks. These authors provide evidence for this bias by showing that fluency can explain short-term stock fluctuations (Alter and Oppenheimer, 2006). Head et al. (2009) confirm this finding and report superior performance of stocks with a more memorable ticker symbol. Green and Jame (2013) find that companies with easy, short names have higher liquidity, higher ownership breadth and higher firm value

However, the focus of this study is not on the general effect of fluency on stock liquidity, firm value or ownership breadth. Instead, it investigates whether the magnitude of the effect differs across industries and firms with various characteristics. This is of interest because previous studies show that behavioral biases are more pronounced in stocks of companies with certain characteristics.

Coval and Moskowitz (1999) document that local equity preference is related to leverage, firm size and output tradability. They report that the home bias tends to be more pronounced in small firms that produce non-traded goods and that have high levels of financial leverage. Moreover, Dahlquist and Robertsson (2001) show that stocks of large and more liquid firms are more likely to be held by foreign investors than those of small firms. This confirms that the magnitude of the home bias and familiarity bias are dependent on firm characteristics. However, little is known about the dependence of the fluency bias on firm characteristics.

Moreover, Cooper et al. (2001) find higher period returns for firms with name changes to dotcom during the Dotcom Bubble. This suggests that the fluency effect was pronounced in the technology industry during the Internet boom. However, not much is known about the differences in the fluency bias across different industries. This paper tries to provide additional information on this topic.

The remainder of this paper is organized as follows. Section (2) discusses the data and methodology used for the assignment of fluency scores. Section (3) examines whether processing fluency has an effect on the full sample. Section (4) investigates whether the fluency bias varies across different industries. Section (5) tests whether the fluency effect is different across different types of firms. Section (6) contains regression with interaction terms to check the results of section (5). Section (6) concludes.

2 Data

2.1 Sample selection

The initial sample includes all securities that are contained in the Center for Research in Security Pricing (CRSP) monthly return file between January 1988 and December 2017. I obtain a list of historical company names and PERMNO security identification numbers, using Wharton Research Data Services. I limit the dataset so it contains common stocks, meaning that the initial sample includes solely stocks that trade with share codes 10 or 11. Moreover, I exclude firms that cannot be assigned to one of the 48 industries classified by Fama and French (1997).

In line with Green and Jame (2013), I remove American Depositary Receipts, Real Estate Investment Trusts and closed-end funds based on share codes (Center for Research in Security Pricing, 2018)[1]. I exclude American Depositary Receipts from the sample, because these are certificates that represent a number of shares of a foreign security (Lang et al., 2003). Since, the aim of this paper is to study the effect of company name fluency on stock returns, it focuses on American firms. Foreign companies are more likely to have names that are non-fluent for American investors. This may bias the results of the tests.

Moreover, I drop closed-end funds from the sample because these are publicly traded investment companies. They are pooled investment funds that contain multiple securities (Lee et al., 1991). It is hard to test the influence of processing fluency on a pool of stocks, since this represents a mixture of different firm names.

Lastly, Real Estate Investment Trusts are companies that solely invest in real estate. These companies have different traits than regular firms and the performance of Real Estate Investment Trusts securities cannot be compared to that of regular stocks (Giliberto, 1990). Therefore, I exclude these stocks.

After meeting all data requirements, the final sample consists of 14,410 firms, 17,851 unique company names and 138,421 firm-years.

2.2 Expansion of historical company names

An important aspect of this study is the method that I use to assign fluency scores to each stock. Previous literature uses roughly two ways to rate stocks based on the processing fluency. The first method is the use of surveys. Through these surveys, participants are, for example, asked to select ticker symbols that they think are clever, cute and memorable (Head et al., 2009). By using the results of these surveys, each ticker symbol is assigned a certain fluency score. However, the use of surveys has found little support among academics whom are active in the field of finance.

The second approach is based on the methodology of Green and Jame (2013), which uses a rational approach to rate stocks based on their fluency. Because of the strength of the measures and the possibility to avoid the use of surveys, I take on the method of Green and Jame (2013) in this study to create fluency

[1] http://www.crsp.com/products/documentation/data-filters

scores. Because this methodology can only be validly used on company names, I neglect ticker symbols.

To calculate fluency scores for companies, I need the full, official name of a firm. I use the cleaned list of historical company names from the final dataset. These names often contain abbreviations, so I expand the historical names and enter them into Microsoft Excel. If an abbreviation is not straightforward enough, I look up the company in the Securities and Exchange Commission Electronic Data-Gathering Analysis and Retrieval (EDGAR) system to get the firm name as reported in the SEC filings.

Following the methodology of Green and Jame (2013), I ignore words in firm names that define the legal name of a company. These specifically include expressions such as Co., Inc, Ltd. and FSB. In addition, I drop the state of incorporation of the company, which is often reported in the name of a bank. Furthermore, I remove hyphens and conjunctions from the company name. So, for example Bay View Federal Savings & Loan Association CA becomes Bay View Federal Savings Loan Association and Word-Wide Technology Inc. is transformed to World Wide Technology. I execute this process for all 17,851 unique historical company names in the final sample

2.3 Length score

After this, I assign fluency scores along three dimensions that give an indication on the ease with which the human brain can process company names. First, short names are easier to understand and remember than longer names (Green and Jame, 2013). Therefore, I count the amount of words a firm's name consists of. This name count is performed on the full company names that are manually created. Based on this count, each word gets a length score assigned. If the adjusted name of a company consists of one word it gets a length score of 3, two words are given a length score of 2 and companies with a greater amount than two words in its name are given a length score of 1.

2.4 Englishness score

The second and third dimensions focuses on the pronounceability of a firm name. Since there may be correlation between past performance, size and pronounceability, I use text-based measure is used to create this dimension. The first is the Englishness of a name. This is based on the linguistic algorithm that is developed by Travers and Olivier (1978). This algorithm states that the Englishness of a n-letter word, which is noted as E', can be estimated using equation (1).

$$E' = -[logF(\#L_1L_2) + \frac{log(F(L_1L_2L_3)}{(F(L_1L_2)} + ... + log[\frac{(F(L_{n-1}L_nL_\#)}{(F(L_{n-1}L_n)}] \tag{1}$$

I estimate the frequency of each three-letter string of each word. This frequency is indicated as $F(L_{k-2}L_{k-1}L_k)$. This is done using data from *The Corpus*

of Contemporary American English. This database provides frequency estimates of English words from over 160,000 texts, ranging from the period 1990 to 2010.

To construct this measure, I divide each word in every company name into three letter strings. I look up the frequency of each string in the *The Corpus of Contemporary American English*. The sum of these string frequencies forms the total frequency for each word. To control for the correlation of word length and total frequency, Englishness is regressed on word length. I use the the residual estimates of this regression as the measure for Englishness. Because one highly non-English word can significantly affect the fluency of a company name, I focus on the word of each name with the lowest frequency. When this information is acquired, I rank the companies ranked on the residuals of the regression. The companies in the bottom quintile of this ranking obtain an Englishness score of 0, all other companies get an Englishness score of 1.

2.5 Dictionary score

The familiarity of a word is related to ease with which it can be pronounced. Green and Jame (2013) propose that a word that can be found in the English dictionary is more likely to be recognizable and familiar than words that are made up by, for example, the firm itself. To assign the dictionary score, I check if each word in the company name passes the Microsoft spell-check in all lowercase letters. If every word in the company name passes the spell-check, the firm gets a dictionary score of 1, all other companies receives a dictionary score of 0.

Based on these three dimensions, I construct an aggregate fluency score which is the sum of the length score, the Englishness score and the dictionary score. I use these aggregate scores to conduct the empirical tests that are tested in the remainder of this paper.

Table 1: Summary statistics on distribution fluency scores

This table contains the summary statistics of the fluency scores of the final sample of this study. The sample consists of all common stocks in the Center for Research in Security Prices and Compustat from 1988 to 2017. *Fluency* scores are the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lowercase letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0.

	Mean	Median	SD
Fluency	3.17	3.00	0.84
Length	2.03	2.00	0.72
Dictionary	0.33	0.00	0.47
Englishness	0.80	1.00	0.40
Observations	138,421	138,421	138,421

Table (1) presents summary statistics regarding the distribution of the Fluency scores. These statistics indicate that the fluency scores of this sample are very similar to those of Green and Jame (2013). The mean total fluency score is 3.17 with standard deviation of 0.84. The mean length score is 2.03 with standard deviation 0.72. The average dictionary score is 0.33 with a standard deviation

0.47. The mean Englishness score is 0.80 with a standard deviation of 0.40. As indicated before, the total amount of firm-years in the sample is 138,421.

Table (2) presents the summary statistics regarding returns and excess returns per fluency score. The companies in the bottom quintile of the categorization have an average return of 1.02% and average excess returns of -1,85%. The category with a fluency score of 2 has average returns and average excess returns of respectively 1.11% and -1.77%. The third fluency category has average returns equal to 1.13% and excess returns equal to -1.72%. The (excess) returns of the fourth category are only marginally larger than those of the first (1.14% and -1.70%). When it comes to average raw returns, the fifth category performs best with an average return of 1.27% and excess return of -1.46%.

The summary statistics show that firms with a higher fluency score perform better than those with a lower fluency scores. However the difference between the categories are only marginally different. The firms in the fifth category obtain 0.25% higher average returns than firms in the first category. A two-sample t-test shows that the difference is statistically significant (t-stat 2.26). The difference in excess returns between the highest and lowest group is 0.39%. This difference is also statistically significant (t-stat 3.19).

The observations per fluency score indicate, however, that the lowest and highest categories are two extremes that do not occur often in the dataset. Because of this, I look at the difference between the mean (excess) returns of the first and second group and the mean (excess) returns of the fourth and fifth group. The difference in average returns between these to groups, even though statistically significant, is nihil. The difference in excess returns between the two highest and two lowest fluency categories is also marginal (0.1%) but statistically significant (t-stat 2.65).

2.6 Other variable construction

After constructing the fluency scores, stock price, shares outstanding, share volume and Standard Industry Classification (SIC) codes are obtained from CRSP. Furthermore, an additional dataset is downloaded from CRSP that contains information on the companies from 1927 to 2017. This additional dataset helps to identify the number of years that a company is listed in the CRSP database. Lastly, the monthly risk-free rate, the Fama and French (1993) risk factors and Momentum factor (Jegadeesh and Titman, 1993) for the 30 years in the sample are downloaded from CRSP.

The same list of PERMNOs is used to acquire firm-level data on long-term debt, debt in current liabilities, stockholders' equity, deferred taxes and investment tax credit, preferred stock (redemption value) from Compustat North America. Since, this paper studies the data on firm-year level and the information from CRSP is monthly and the data obtained from Compustat is quarterly, the annual averages of all variables are calculated. For each firm-year, additional control variables are computed. A complete list of these variables and their definitions is included in Appendix B.

In line with Green and Jame (2013), I winsorize all firm-level variables at the

1st and 99th percentiles to eradicate the effect of outliers. Moreover, I remove all observations with returns that lie outside the 1st and 99th percentile.

Table 2: Return and excess returns statistics per fluency score

This table contains the summary statistics of return and excess return per fluency score. The sample consists of all common stocks in the Center for Research in Security Prices and Compustat from 1988 to 2017. Returns are the annual average of monthly holding period returns acquired from CRSP. Excess returns are the annaual average of monthly holding period returns minus the yield of a 1-month Treasury bill. *Fluency* scores are the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0.

Fluency	Returns	Excess Returns
Score 1		
Mean	0.0102	-0.0185
Median	0.0102	-0.0181
SD	0.0446	0.0496
Observations	2,461	2,461
Score 2		
Mean	0.0111	-0.0177
Median	0.0110	-0.0178
SD	0.0463	0.0508
Observations	27,449	27,449
Score 3		
Mean	0.0113	-0.0172
Median	0.0111	-0.0171
SD	0.0462	0.0508
Observations	57,580	57580
Score 4		
Mean	0.0114	-0.0170
Median	0.0111	-0.0168
SD	0.0473	0.0518
Observations	46,103	46,103
Score 5		
Mean	0.0127	-0.0146
Median	0.0118	-0.0155
SD	0.0447	0.0493
Observations	4,828	4,828

Table (3) presents the cross-sectional summary statistics for the final sample from 1988 to 2017. The average firm is around 15 years old, has a market capitalization of about 1.8 billion US dollars, a book-to-market ratio of 69%, a book leverage ratio of 34% and a turnover ratio of 122%. Since, the median of size is significantly different from the mean and the standard deviation is large, this variable is log-transformed to eliminate the potential effect of outliers. The variables age and book-to-market are also log-transformed as the large standard deviations of these variables indicate that outliers may affect the results of the regression analysis.

Table 3: Summary Statistics

This table reports the cross-sectional summary statistics of the final sample used for this study. The sample consists of all common stocks in the Center for Research in Security Prices and Compustat from 1988 to 2017. *Size* is defined as market capitalization, calculated as share price times shares outstanding. *Age* is the number of years a firm is available in the CRSP. *Book-to-market ratio* is the book value of equity divided by the market value of equity, where book value of equity is defined as the value of stockholders' equity plus deferred taxes and investment credit minus the redemption value of preferred stock. *Book leverage* is total debt divided by the sum of total debt and book equity. *Turnover* is the average annual turnover scaled by shares outstanding. Share price, average turnover and shares outstanding are obtained from CRSP. The value of stockholders' equity, deferred taxes and investment credit, redemption value of stock and total debt are obtained from Compustat North America.

	Mean	Median	SD
Size	1,818,358	169,283	5,766,970
Age	14.81	10	15.22
B/M	0.69	0.55	1.62
Leverage	0.34	0.30	0.32
Turnover	1.22	0.77	1.32
Intangible assets	0.16	0.09	0.18
Observations	138,421	138,421	138,421

3 The effect of fluency on returns

3.1 The effect of company name fluency on excess returns

This part of the study tests whether there is a relation between fluency and excess returns. In order to do this, I estimate equation (2).

$$Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \epsilon_{it} \tag{2}$$

The test described above is a panel regression that does not include control variables. $Fluency_{i,t-1}$ is a categorical variable that ranges from 1 to 5. This variable is lagged, because an investor observes a company name in period $t-1$ and makes his investment decision in period t. Therefore, the test is performed as a lead-lag regression. The only variable of interest is $Fluency_{i,t-1}$.

After this simple regression, I test whether fluency affects returns after controlling for firm characteristics that are possible determinants of returns. In order to study this, I estimate equation (3).

$$Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 \mathbf{X_{i,t-1}} + \epsilon_{it} \tag{3}$$

$\mathbf{X_{i,t-1}}$ is a vector of firm characteristics. Using this vector, I control for the effect of firm size, book-to-market ratio, historical returns, share turnover and book leverage is controlled. I leave out firm age as control variable, because, due to its deterministic nature, there are econometric issues concerning the variable.

To control for time-invariant firm characteristics, I add firm-fixed effects. This helps to focus on within-firm changes in the fluency effect. In addition, I add HAC Newey-West standard errors to control for heteroskedasticity and autocorrelation.

Every test I perform in this study includes these firm-fixed effects and robust standard errors.

If the effect of fluency remains economically and statistically significant, the differences between fluent and non-fluent firms cannot solely be explained by cross-sectional differences in characteristics.

Table (4) reports the results of the panel regressions of excess returns on fluency and other firm characteristics. Model (1) presents the results of the estimation of equation (2). The coefficients of this model indicate that there is a significant and positive relation between excess returns and fluency. A change from the lowest fluency category to the highest category results in a return increase of 108 basis points. This means that a security, solely because it has a fluent name, performs significantly better than stocks with non-fluent names.

Table 4: Company name fluency and returns

This table reports the estimates from panel regression of excess returns on fluency and other firm characteristics. The first model shows the results of a simple panel regression of excess returns on fluency. The second model reports the estimates of a multiple panel regression of excess returns on fluency and other firm characteristics that are known determinants of returns. Excess returns are obtained from CRSP. *Fluency* scores are the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Definitions of the control variables are reported in Appendix B.

	(1) Excess Returns	(2) Excess Returns
Fluency	0.0022	0.0021
	(3.67)	(2.50)
Log(Size)		-0.0038
		(-12.40)
Log(B/M)		0.0113
		(26.79)
Book leverage		0.0170
		(12.03)
Return		-0.1079
		(-27.68)
Turnover		-0.0000
		(-0.14)
Fixed effects	Yes	Yes
N	138,421	110,182

Model (2) also estimates the influence of fluency on excess returns. However, this model controls for the effect of several determinants of returns. The results show that the fluency effect is still both statistically and economically significant. Adding the control variables reduces the magnitude of the effect slightly, but the economic difference between two fluency categories is still 21 basis points. Therefore, fluency has a positive effect on excess returns. These findings are in line with those of, for example, Alter and Oppenheimer (2006) and Head et al. (2009).

4 Variation of the fluency effect across industries

In order to test whether the influence of processing fluency on returns is more pronounced in certain industry, I divide the dataset into multiple subsamples. I categorize the firms based on the Fama and French (1997) industry classification. Based on four-digit SIC-codes, I assign firms to onesigned to one of the 48 industries that Fama and French (1997) identify.

After categorizing the firms in the sample, equation (4) is estimated on each of the 48 subsamples. By looking at the differences between the estimated betas, the effect of processing fluency can be compared across the various industries.

$$Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \epsilon_{it} \tag{4}$$

The results of the tests are included in Appendix A. The estimates indicate that the influence of fluency depends on industry and does not affect companies in all industries equally. When no control variables are added to the model, the fluency effect is statistically significant for 9 different industries, namely the tobacco, fabricated metal products, automobile, ship building, container, mining, utility, banking and catering industries. Interestingly, the fluency effect is negative for mining sector and the sector that fabricates products from metal. This is not in line with expectations, since the fluency effect has most commonly a positive influence on returns.

For the industries in which the fluency effect is positive and statistically significant, there are some differences in magnitude between the different subsamples. For the banking industry, the effect may be statistically present, however economically speaking, the coefficient is marginal. This is the same for the restaurant, hotel and motel sector. The fluency effect is most pronounced in the tobacco, ship building, automobile and container industries. These are, generally speaking, industrial companies that are dependent on fixed assets, like property, plant and equipment, and working capital.

In order to check whether the previous results are robust to the influence of firm-level variables that determine returns, a vector of firm-characteristics is added to the model. The vector consists of the same firm characteristics that are used in section (3). In line with the methodology of paragraph (3.1), the equation (5) is estimated.

$$Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 \mathbf{X_{i,t-1}} + \epsilon_i t \tag{5}$$

The regression is performed for all 48 subsamples. Again, the coefficient on the categorical variable fluency is of interest. The various coefficients of fluency that result from the multiple regressions are interpreted to compare the fluency effect across the different industries. Moreover, the coefficients and t-statistics of this test are compared to the results presented in table (17) to see whether the influence of processing fluency on excess returns is affected by adding the control variables.

The regressions including control variables show that the fluency coefficients of some industries that are significant in the previous test become statistically insignificant. This is the case for the mining and banking industry. The fluency effect remains present in the tobacco, fabricated metal products, utility and container sectors. Moreover, after adding the control variables, fluency still affects the returns of stocks in the restaurant, hotel and motel industry. Generally speaking, the results infer that the fluency bias is most pronounced in the industrial sectors.

5 Variation of the fluency effect across types of firms

In this section, I study whether the fluency effect is more pronounced in firms with certain characteristics. In order to study the differences between the influence of processing fluency across firm types, I divide the companies into quintiles based on a certain characteristic. Equations (2) and (3) are estimated for each of the quintiles. The fluency coefficients of the various quintiles are then compared to see whether the effect differs across the quintiles. This way, the strength of the fluency effect is compared between firms with different firm characteristics.

The characteristics that are investigated are firm size (market capitalization), book-to-market ratio, firm age, asset liquidity and book leverage ratio. The tables containing the results of the relation between the fluency effect and these two firm characteristics are left out of this study. The tests on these two firm characteristics show there is no evident relation between the behavioral bias and asset liquidity and book leverage. So, the results indicate that there is no difference in fluency effect between firms with different levels of asset liquidity and book leverage. Because of this, I do not report these tables.

5.1 Fluency effect across small and large firms

Table 5: Fluency effect across smaller and larger firms without control variables

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their size (market capitalization). The first column presents the results of the quintiles containing the smallest firms and the fifth column shows the estimates of the quintile containing the largest firms. Excess returns are obtained from CRSP. *Fluency* scores are the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0

	Size 1	Size 2	Size 3	Size 4	Size 5
Fluency	0.0048	0.0022	-0.0006	0.0029	0.0028
	(3.09)	(1.35)	(-0.39)	(2.12)	(2.45)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	27,684	27,684	27,684	27,685	27,684

Table 6: Fluency effect across smaller and larger firms with control variables

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their size (market capitalization). The first column presents the results of the quintiles containing the smallest firms and the fifth column shows the estimates of the quintile containing the largest firms. Excess returns are obtained from CRSP. *Fluency* is the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Descriptions of the control variables are presented in Appendix B.

	Size 1	Size 2	Size 3	Size 4	Size 5
Fluency	0.0049	0.0025	0.0008	0.0039	0.0022
	(1.72)	(1.17)	(0.35)	(2.20)	(1.65)
Log(Size)	-0.0151	-0.0337	-0.0427	-0.0355	-0.0065
	(-14.35)	(-25.18)	(-31.63)	(-37.85)	(-9.76)
Log(B/M)	0.0048	0.0036	0.0048	0.0127	0.0113
	(4.09)	(3.38)	(4.52)	(13.65)	(13.05)
Leverage	0.0072	0.0080	0.0133	0.0178	0.0337
	(1.82)	(2.29)	(4.02)	(6.10)	(12.47)
Return	-0.1750	-0.2513	-0.2635	-0.2509	-0.1404
	(-17.12)	(-27.09)	(-29.70)	(-29.02)	(-15.23)
Turnover	-0.0057	-0.0055	-0.0016	0.0014	0.0055
	(-6.83)	(-8.69)	(-2.83)	(2.95)	(13.73)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	16,429	21,605	22,708	24,094	25,346

Table (5) contains the results of the simple regressions of excess returns on fluency for quintiles containing firms with different sizes. The first column presents the coefficients of the regression on the smallest firms in the sample and the fifth model shows the results of the quintile with the largest firms.

The estimates show that the fluency effect is not statistically significant for all subsamples. The effect is only significant for the subsample containing the smallest firms and the two subsamples containing the largest firms. The effect is marginally signficant for the firms in the second quintile. The magnitude of the fluency effect is greatest for smallest firms. An increase in fluency level results in a gain of excess returns of 48 basis points per year. The strength of the effect is smallest for the second quintile of firms. The estimation for the fourth and fifth category are roughly the same.

The result that the fluency effect is most pronounced in the smallest firms is in line with expectations. Previous literature suggests that more information is available for larger firms than for smaller firms (Atiase, 1985) and (Grant, 1980). Collins et al. (1987) indicate that traders and analyst process the information of larger firms more intensively than that of smaller firms. This higher information efficiency for larger firms should result in less mispricing. Therefore, the fluency effect should be less exhibited in firms with a higher market capitalization.

However, the results on the other categories do not point as strongly to the relation that the fluency effect is decreasing in size. Statistically, the result is weaker for the second and third quintiles and stronger for the fourth and fifth

quintiles. The difference between the categories containing the largest firms is small.

The results may be affected by a familiarity bias, since larger firms are usually better known to investors than smaller firms. The lack of control variables may provide a possibility for the familiarity bias to take over and bias the results slightly.

Table (6) shows the estimates of the quintile regressions, controlling for firm characteristics. The results of this test change slightly compared to those of the model without any control variables. The effect is only significant for the fourth quintile. The coefficients are marginally significant for the first, second and fifth quintile. The economic magnitude of the fluency effect is still greatest for the quintile containing the largest firms in the sample. When looking at the results of the fourth and fifth categories, the fluency effect decreases substantially. Overall, the results suggest that the fluency effect is decreasing in firm size.

This supports the reasoning that in the previous test the lack of controls allows a familiarity bias to bias the results slightly. The difference between the coefficients of the fourth and fifth quintiles becomes significantly larger. Since, the firms in the fifth quintile are most likely the best known firms, the results of table (5) on these categories may be influenced by the familiarity bias.

The finding that the fluency effect decreases in size is in line with expectations. As indicated before, large firms experience increased analyst coverage, which should result in less mispricing. This reasoning is supported by the results of table (6). So, when controlling for, among other, asset liquidity and book-to-market, the fluency effect is more pronounced in smaller firms than in larger firms.

5.2 Fluency effect across firms with a high and low book-to-market ratio

Table 7: Fluency effect across firms with a high and low B/M ratios without control variables

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their book-to-market ratio. The first column presents the results of the quintiles containing the firms with the lowest B/M ratio and the fifth column shows the estimates of the quintile containing the firms with the hight B/M ratio. Excess returns are obtained from CRSP. *Fluency* is the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0.

	B/M 1	B/M 2	B/M 3	B/M 4	B/M 5
Fluency	-0.0014	0.0030	0.0019	0.0019	0.0050
	(-0.81)	(2.07)	(1.34)	(1.59)	(3.71)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	27,209	27,209	27,210	27,210	29,583

Table (7) presents the estimates of the regressions on the different quintiles. The results show that fluency only has an effect on the firms in the second, fourth and fifth quintiles. The fluency effect is most pronounced in firms with the highest book-to-market ratio. The effect is least strong for the companies in the fourth quintile. Therefore, even though the results are weak and slightly ambiguous, table (7) shows that firms with a higher book-to-market ratio, generally exhibit a lower fluency effect.

This result is not in line with expectations. Firms that have a market value that is much larger than their book value are often mispriced. Therefore, a low book-to-market ratio (or high market-to-book ratio) is often considered to be a sign of mispricing (Rosenberg et al., 1998) and (Lakonishok et al., 1994). Because of this, the expectation is that the fluency effect is more pronouned in firms with a low book-to-market ratio. However, the results of this test show the contrary.

Table 8: Fluency effect across firms with a high and low B/M ratios with control variables

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their book-to-market ratio. The first column presents the results of the quintiles containing the firms with the lowest B/M ratio and the fifth column shows the estimates of the quintile containing the firms with the hight B/M ratio. Excess returns are obtained from CRSP. *Fluency* is the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Descriptions of the control variables are presented in Appendix B.

	B/M 1	B/M 2	B/M 3	B/M 4	B/M 5
Fluency	-0.0011	0.0018	0.0019	0.0032	0.0012
	(-0.39)	(0.99)	(1.19)	(2.22)	(0.61)
Log(Size)	-0.0111	-0.0057	-0.0027	0.0025	0.0011
	(-13.19)	(-9.26)	(-4.31)	(3.84)	(1.26)
Log(B/M)	0.0042	0.0269	0.0404	0.0463	0.0315
	(4.38)	(21.56)	(27.50)	(33.74)	(23.98)
Leverage	0.0116	0.0367	0.0272	0.0148	0.0003
	(3.22)	(12.11)	(8.47)	(4.35)	(0.07)
Return	-0.1512	-0.1993	-0.2202	-0.2402	-0.2065
	(-13.93)	(-20.32)	(-20.04)	(-21.77)	(-20.26)
Turnover	-0.0028	0.0003	0.0019	0.0022	0.0008
	(-4.60)	(0.72)	(3.49)	(4.27)	(1.14)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	16,860	23,198	23,923	23,832	22,369

Table (8) contains the results of the regressions that control for firm characteristics that may have an effect on excess returns. When adding the control variables, the effect of fluency is eliminated for the second and fifth quintiles of firms. The effect remains only significant for the firms in the fourth quintile. Even though, the results are statistically not strong, it seems that there may be a relation between book-to-market ratio and the fluency effect. This relation may not be captured by this empirical test. Therefore, I investigate this relationship further in the next section.

5.3 Fluency effect across older and younger firms

Table 9: Fluency effect across older and younger firms without control variables

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their age. The first column presents the results of the quintiles containing the youngest firms and the fifth column shows the estimates of the quintile containing the oldest firms. Excess returns are obtained from CRSP. *Fluency* is the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0.

	Age 1	Age 2	Age 3	Age 4	Age 5
Fluency	-0.0021	0.0022	-0.0002	0.0006	0.0034
	(-0.52)	(0.87)	(-0.14)	(0.35)	(2.72)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	23,497	27,777	31,453	29,359	26,335

The results presented in table (9) indicate that the fluency effect is only economically and statistically significant for the oldest firms in the sample. The coefficients of the other quintiles are not statistically significant, indicating that these subsamples most likely do not exhibit the fluency effect. However, fluency has a substantial effect on the 20% oldest firms in the sample. An increase in fluency level, results in an increase in excess returns of 34 basis points per year.

After adding control variables to the regression, this result is not eliminated. The fluency effect is still only significant for the quintile containing the oldest firms in the sample. The magnitude of the effect decreases marginally from 32 basis points per fluency level per year to 30 basis points. However, both results are economically significant. These results again suggest that the fluency effect is only present in the oldest firms in the sample.

These results are not in line with expectations. It is expected that the fluency effect is more pronounced in younger firms than in older firms. Calantone et al. (2002) indicates that the market information on older firms is more efficient than on younger. Due to more effective supply of market information, investors are able to make better informed decisions when it comes to older firms. This results in less mispricing. Therefore, a behavioral bias, such as the fluency effect is less pronounced in older than in younger firms. However, the estimates of table (9) and (10) contradict this line of reasoning.

A possible explanation for this finding is the familiarity bias. The oldest firms in the sample have been on the market for the longest period of time. Whereas, the firms in the first four quintiles range from 0 to 24 years old, firms in the fifth sample are between 25 and 92 years old. So, these are companies that have been traded in the stock market for many years, which may cause them to be well known to investors. It is possible that this results in a familiarity bias that is not present in the other stocks.

Table 10: Fluency effect across older and younger firms with control variables

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their age. The first column presents the results of the quintiles containing the youngest firms and the fifth column shows the estimates of the quintile containing the oldest firms. Excess returns are obtained from CRSP. *Fluency* is the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lowercase letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Descriptions of the control variables are presented in Appendix B.

	Age 1	Age 2	Age 3	Age 4	Age 5
Fluency	0.0000	-0.0005	0.0000	0.0008	0.0030
	(0.59)	(-0.12)	(0.01)	(0.38)	(2.08)
Log(Size)	-0.0478	-0.0330	-0.0250	-0.0144	-0.0020
	(-18.49)	(-22.89)	(-25.20)	(-16.52)	(-2.58)
Log(B/M)	-0.0087	-0.0003	0.0012	0.0096	0.0110
	(-3.79)	(-0.23)	(1.16)	(9.40)	(10.71)
Leverage	0.0007	0.0052	0.0099	0.0101	0.0224
	(0.10)	(1.12)	(2.81)	(3.08)	(7.19)
Return	-0.2986	-0.2614	-0.1734	-0.1333	-0.1320
	(-26.42)	(-29.88)	(-22.12)	(-15.92)	(-13.97)
Turnover	-0.0044	-0.0033	-0.0021	-0.0008	0.0066
	(-5.40)	(-4.89)	(-4.15)	(-1.55)	(11.89)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	12185.0000	22018.0000	26477.0000	25378.0000	24124.0000

However, by adding the control variables of the next test, a possible familiarity bias may be eliminated from the results. If the estimates follow the same pattern in the estimates with control variables, it is more likely that the result is due to the fluency effect and not due to the familiarity bias. Because the effect of fluency does not change after adding control variables, it is most likely that the results of table (9) and table (10) are an exhibition of the fluency effect instead of the familiarity bias.

Another option is that the results are driven by the survivorship bias. The survivor bias is the tendency of investors to overestimate historical performance of a stock. Because survivors perform sufficiently to exist for a longer period of time, they are expected to be better. Moreover, previous literature finds that the survivor bias may lead to other behavioral biases in cross-sectional performance (Brown et al., 1992). Since all firms in the fifth quintile are "survivors", it is possible that the survivorship bias may enhance the existence of other behavioral biases such as the fluency effect.

However, this result still goes against the most logical line of reasoning.

The results of this test may not be in line with the reasoning based on efficient information distribution of older firms. However, they can also be explained by the survivorship bias making it possible for the fluency effect to exist in the quintile containing the oldest firms.

5.4 Fluency across firms with low and high levels of intangible assets

Table 11: Fluency effect across firms with low and high levels of intangible assets

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their intangible assets. The first column presents the results of the quintiles containing the firms with the lowest amount of intangible assets relative to their total assets. The fifth column presents the results of the quintiles containing the firms with the highest amount of intangible assets relative to their total assets. Excess returns are obtained from CRSP. *Fluency* is the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0.

	Intangibles 1	Intangibles 2	Intangibles 3	Intangibles 4	Intangibles 5
Fluency	0.0006	-0.0002	0.0000	0.0014	0.0023
	(0.95)	(-0.29)	(0.01)	(2.17)	(3.36)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	16,898	16,898	16,898	16,898	16,899

Table (11) reports the estimates of the regression of fluency on excess returns per quintile. The estimates show that the fluency effect is only statistically significant for the fourth and fifth quintile. The coefficients of the other quintiles do not hold any statistical power. In the fourth category, a jump to a higher fluency level increases excess returns with 13 basis points per year. An increase in fluency with one level raises excess returns with 23 basis points per year in the fifth quintile. This indicates that the fluency effect is most pronounced in firms with higher levels of intangible assets.

It is expected that the fluency effect is stronger for firms with higher levels of intangible assets. Barth et al. (2001) indicate that the values of intangible assets are often not disclosed, analyst coverage of companies with more intangible assets are likely to be less informative. Because of this, the information disclosed on firms with more intangible assets is less valuable. This lower information efficiency on firms with a lot of intangible asset may cause these firms to be more often mispriced than those with less intangible assets. Therefore, the results of table (11) are in line with expectations.

Table (12) shows the results of the regression with control variables. The coefficient estimates do not change after controlling for various firm characteristics. The fluency effect is still only significant for the firms in the fourth and fifth quintile. The economic magnitude of the coefficients do not change significantly.

So, even after adding control variables to the regression, the fluency effect is not exhibited in three quintiles containing firms with the least intangible assets. The fluency effect does exist in the categories with the companies the highest amount of intangible assets. For these categories, the bias is the strongest for the fifth quintile. This is in line with reasoning based on the availability of information on firms with more intangible assets.

Table 12: Fluency effect across older and younger firms with control variables

This table presents the estimates of regressions of excess returns on fluency. Firms are categorized into quintiles based on their intangible assets. The first column presents the results of the quintiles containing the firms with the lowest amount of intangible assets relative to their total assets. The fifth column presents the results of the quintiles containing the firms with the highest amount of intangible assets relative to their total assets. Excess returns are obtained from CRSP. *Fluency* is the aggregate of length, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Descriptions of the control variables are presented in Appendix B.

	Intangibles 1	Intangibles 2	Intangibles 3	Intangibles 4	Intangibles 5
Fluency	-0.0000	0.0004	0.0001	0.0013	0.0024
	(-0.03)	(0.61)	(0.19)	(1.89)	(3.42)
Log(Size)	0.0012	0.0004	0.0002	0.0010	-0.0000
	(3.79)	(1.53)	(0.83)	(3.38)	(-0.09)
Log(B/M)	0.0119	0.0084	0.0078	0.0096	0.0082
	(16.87)	(11.27)	(10.68)	(12.24)	(10.53)
Leverage	-0.0096	-0.0017	-0.0002	0.0077	0.0135
	(-4.62)	(-0.76)	(-0.10)	(3.29)	(5.62)
Return	-0.0692	-0.0989	-0.1035	-0.1263	-0.1183
	(-6.12)	(-9.23)	(-10.20)	(-11.80)	(-11.86)
Turnover	0.0006	0.0019	0.0009	0.0012	-0.0008
	(1.17)	(4.13)	(1.87)	(2.60)	(-1.62)
Fixed effects	Yes	Yes	Yes	Yes	Yes
N	14,104	14,296	14,343	14,464	14,247

6 Interaction between fluency and firm characteristics

The previous section shows that a relation exists between the fluency effect and various firm characteristics, namely size, age and (most likely) book-to-market ratio. To check the robustness of these tests, I perform additional regressions that contain interaction terms between fluency and these variables. I do this to test whether the magnitude of the fluency effect depends on any of these firm characteristics. Statistically and economically significant interaction terms indicate that there is indeed a relation between the fluency effect and the variables.

6.1 Interaction between fluency and firm size

Subsection (5.1) shows that the fluency effect decreases in size of the firm. In this section, I perform two additional regressions to check the robustness of these results. I do this by testing whether the magnitude of the fluency effect depends on firm size. I estimate the equation (6) without any control variables to see the interaction between the two variables of interest.

$$\begin{aligned} Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 Size + \\ \beta_3 Fluency_{i,t-1} * Size_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{6}$$

To see whether the results from the estimation of equation (6) hold when adding control variables, the same vectors as in the previous sections is added to the regression. These results are estimated using equation (7).

$$\begin{aligned} Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 Size + \\ \beta_3 Fluency_{i,t-1} * Size_{i,t-1} \beta_4 \mathbf{X_{i,t-1}} + \epsilon_i i,t \end{aligned} \tag{7}$$

The results of this regression are presented in table (13). The estimates of model (1) indicate that the stand-alone effect of fluency is an increase in excess returns of 3.8% per fluency score. The magnitude of this estimation is much greater than that of model (1) of table (4).

The interaction effect between fluency and size is also statistically significant. The estimates indicate that an increase inside decreases the fluency effect significantly. A one percent increase in size results in a 0.000028 percent decrease in the fluency effect. This result seems marginal. However, firm size is a variable with a large standard deviation, indicating that the values of this variable are widely spread. Therefore, size can significantly reduce the fluency effect. This in line with the results of table (6).

When adding control variables to the regression, the magnitude of the variables of interest decreases. The stand-alone effect of fluency lowers to 2.3% per category increase in fluency. The interaction effect also becomes less strong. A one percent increase in size results in a 0.000017 percent decrease in the fluency effect. Again, this may seem marginal. But since the data on size is widely spread, size may reduce the fluency effect significantly.

These results are also in line with those of table (6). Therefore, this test confirms that the fluency effect is less pronounced in larger firms. This in line with intuition since mispricing should be most exhibited in smaller firms due to less trader and analyst coverage (Collins et al., 1987).

Table 13: OLS with interaction terms between fluency and firm size

This table presents the estimates of regressions of excess returns on fluency, including interactions terms between fluency and firm size. *Fluency* scores are the aggregate of *length*, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Definitions of size and the other control variables are included in Appendix B.

	(1) Excess Returns	(2) Excess Returns
Fluency	0.0382	0.0230
	(27.67)	(14.64)
Log(Size)	0.0000	0.0000
	(27.30)	(24.40)
Fluency × Log(Size)	-0.0028	-0.0017
	(-32.78)	(-16.77)
Log(B/M)		0.0110
		(26.99)
Leverage		0.0159
		(11.36)
Return		-0.1123
		(-28.76)
Turnover		0.0001
		(0.62)
Fixed effects	Yes	Yes
N	115,728	110,182

6.2 Interaction between book-to-market ratio and fluency

Subsection (5.2) shows that the relation between book-to-market ratio and the fluency effect is ambiguous. However, the results show some indication that there may be a relation between the two variables. In order to see whether this is the case, I estimate equation (**??**) to see whether there is an interaction effect between the book-to-market ratio and the fluency effect.

$$\begin{aligned} Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 B/M + \\ \beta_3 Fluency_{i,t-1} * B/M_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{8}$$

After this, I add a vector of firm characteristics to check whether the results of equation (8) hold after adding control variables. In order to do this, I estimate equation (9).

$$\begin{aligned} Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 B/M_{i,t-1} + \\ \beta_3 Fluency_{i,t-1} * B/M_{i,t-1} + \beta_4 \mathbf{X_{i,t-1}} + \epsilon_{i,t} \end{aligned} \tag{9}$$

Table 14: OLS with interaction terms between fluency and book-to-market ratio

This table presents the estimates of regressions of excess returns on fluency, including interactions terms between fluency and book-to-market ratio. *Fluency* scores are the aggregate of *length*, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Definitions of book-to-market and the other control variables are included in Appendix B.

	Excess Returns	Excess Returns
Fluency	0.0018	0.0032
	(2.82)	(3.66)
B/M	0.0007	0.0003
	(2.04)	(0.53)
Fluency × B/M	-0.0002	-0.0004
	(-1.96)	(-2.30)
Log(Size)		-0.0083
		(-30.86)
Leverage		0.0059
		(5.41)
Return		-0.1268
		(-33.81)
Turnover		0.0002
		(0.95)
Fixed effects	Yes	Yes
N	136,047	113,489

The results of table (14) show that the stand-alone effect of fluency is statistically significant. The fluency effect results in an increase in excess returns of 18 basis points per fluency level per year. This effect is slightly lower than that reported in table (4). Moreover, the interaction coefficient between between fluency and book-to-market is also statistically significant and indicates that the strength of the fluency effect decreases when the book-to-market ratio of a firm becomes larger.

These results are in line with the expectation that the fluency effect is more pronounced in firms with a low book-to-market ratio. As expressed in the previous subsection, a low book-to-market ratio is an indicator of mispricing. Therefore, a mispricing, such as the fluency effect, is expected to be stronger in firms with a low book-to-market ratio. The results of model (1) of table (14) are in line with this reasoning. The results are not consistent with those of table (7) and (8). Those results were ambiguous and statistically weak. The estimation of model (1) of table (14) are stronger.

After adding control variables to the regression, the estimates of the coefficients change slightly. The stand-alone effect of fluency increases to 24 basis points per fluency level, whereas the interaction coefficient decreases to -5 basis points. However, the overall conclusion of this test does not change after controlling for various firm characteristics.

The main conclusion, I derive from these two regressions is that that the flu-

ency effect is decreasing in book-to-market ratio. This indicates that the returns of stocks of firms with a high market value relative to their book value are more likely to be affected by fluency.

6.3 Interaction between fluency and firm age

Subsection (5.3) shows that the fluency effect is more pronounced in older than in younger firms. In order to check the robustness of this test, I also estimate a regression with interaction terms between fluency and firm age. Equations (10) and (11) are estimated to see whether the strength fluency effects depends on the age of a firm.

$$\begin{aligned} Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 Age_{i,t-1} + \\ \beta_3 Fluency_{i,t-1} * Age_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{10}$$

$$\begin{aligned} Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 Age_{i,t-1} + \\ \beta_3 Fluency_{i,t-1} * Age_{i,t-1} + \beta_4 \mathbf{X_{i,t-1}} + \epsilon_{i,t} \end{aligned} \tag{11}$$

Table (15) contains the results of the regression including an interaction term between fluency and age. The estimates of model (1) indicate that the stand-alone effect of fluency is negative. However, the interaction coefficient shows that there is an interaction between age and fluency. A one year increase in age results in an increase in fluency effect of 0.0004. The youngest firm in the sample is 0 while the oldest firm is 94. Because of this wide spread, age may have a significant influence on the the fluency effect. This is in line with the results of table (9) and table (10).

The full specification shows that the estimates of the coefficients change slightly after adding control variables. The stand-alone of fluency is still negative but has decreased to -0.0162 per fluency level per year. The coefficient on the interaction between age and fluency increases slightly from 0.0004 to 0.0007. This means that if the firm becomes one year older, the fluency effect increases with 0.0007. So, overall adding control variables to the regression does not change the conclusion of the previous paragraph.

The interaction effect between fluency and age indicate that the (positive) fluency effect is stronger for older firms than for younger firms. This in line with the findings of subsection (5.3). Therefore, this test confirm the finding of that subsection.

Table 15: OLS with interaction terms between fluency and firm age

This table presents the estimates of regressions of excess returns on fluency, including interactions terms between fluency and firm age. *Fluency* scores are the aggregate of *length*, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Definitions of age and the other control variables are included in Appendix B.

	(1) Excess Returns	(2) Excess Returns
Fluency	-0.0085	-0.0162
	(-12.91)	(-19.26)
Age	0.0036	0.0058
	(9.92)	(12.20)
Fluency × Age	0.0004	0.0007
	(42.83)	(45.56)
Log(Size)		-0.0156
		(-43.38)
Log(B/M)		0.0017
		(4.15)
Leverage		-0.0011
		(-0.82)
Return		-0.1228
		(-30.10)
Turnover		-0.0010
		(-4.49)
Fixed effects	Yes	Yes
N	114,846	104,807

6.4 Interaction between fluency and intangible assets

The results of table (11) and (12) show that the fluency effect is decreasing when intangible assets increase. In this section, I add interaction terms between these two variables to double-check the relation between fluency and intangible assets. I do this by estimating equations (12) and (13).

$$Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 Intangibles_{i,t-1} + \beta_3 Fluency_{i,t-1} * Intangibles_{i,t-1} + \epsilon_{i,t} \tag{12}$$

$$Return_{i,t} - Rf_t = \alpha_0 + \beta_1 Fluency_{i,t-1} + \beta_2 IntangibleAssets_{i,t-1} + \beta_3 Fluency_{i,t-1} * Intangibles_{i,t-1} + \beta_4 \mathbf{X_{i,t-1}} + \epsilon_{i,t} \tag{13}$$

Table (16) contains the results of the regressions including an interaction term between fluency and intangible assets. Model (1) shows that the effect of fluency by itself is not statistically significant. However, the coefficient of interest in this case is the interaction term between fluency and intangible assets.

Looking at the, t-statistic the interaction term is statistically significant. The coefficient shows that the relation between fluency and intangible assets is positive. When intangible assets become larger, the fluency effect increases in magnitude. This indicates that there is a positive relation between fluency and intangible assets. Model (1) therefore confirms the findings of subsection (5.4).

Model (2) of table (16) shows that the results of model (1) do not change significantly after adding control variables to the regression. The coefficient of fluency itself is still statistically insignificant. The interaction term between fluency and intangible assets is, as in model (1), statistically significant. This means that there is a positive relation between the two variables of interest after adding the controls.

This regression confirms that the fluency effect varies across firms with different levels of intangible assets. This is in line with the findings of subsection (5.4) and the informational reasoning that is proposed in that subsection. Because the information provision of firms with more intangible assets is weaker than that of firms with less intangible assets, mispricings are more likely to occur in the stocks of the first type of firms (Barth et al., 2001). The findings of these regressions are in line with this reasoning.

Table 16: OLS with interaction terms between fluency and intangible assets

This table presents the estimates of regressions of excess returns on fluency, including interactions terms between fluency and tangible assets. *Fluency* scores are the aggregate of *length*, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. Definitions of intangible assets and the other control variables are included in Appendix B.

	Excess Returns	Excess Returns
Fluency	-0.0009	-0.0007
	(-0.98)	(-0.61)
Intangible assets	-0.0221	-0.0097
	(-2.98)	(-1.09)
fluency × Intangible assets	0.0085	0.0072
	(3.92)	(2.77)
Log(Size)		-0.0058
		(-13.49)
Log(B/M)		0.0113
		(19.99)
Leverage		0.0164
		(9.22)
Return		-0.1287
		(-26.32)
Turnover		0.0004
		(1.53)
Fixed effects	Yes	Yes
N	84,490	71,453

7 Conclusion

Recent literature shows that psychological biases have an influence on the decisions of investors. One example is the fluency bias. This bias states that stocks that are easier to process are better liked by investors. Previous studies show that fluency has an effect on the stock market. Stocks that have fluent names tend to outperform stocks that have names that are less easy to process. This study focuses on the effect of fluency on stock returns. In particular, it looks at differences of the effect between different industries and firms with different characteristics.

This paper finds that there are differences in the fluency effect between different industries. For most industries, the fluency of a company name does not have an effect on excess returns. However, company names do have an influence on the stock returns of firms active in the tobacco, fabricated metal products, utility and container sectors. The fluency effect is also more pronounced in companies that operate in the restaurant, hotel and motel industry.

The results also show that the fluency effect does not vary across firms with different levels of asset liquidity and book leverage. Moreover, the results indicate fluency differs across firms with other characteristics. The empirical tests do suggest that the fluency effect is more pronounced in older firms, firms with a high market capitalization, a low book-to-market ratio and a high level of intangible assets. characteristics. Stock returns of older firms, firms with a high market capitalization, a low book-to-market ratio, and a high level of intangible assets. These results are supported by the regressions containing interaction terms between fluency and the characteristics of interest.

The findings of this paper suggest that investors and other participants of the stock market are able to use an investment strategy that makes use of the fluency effect. However, this study shows that exploiting the fluency effect is not as straightforward as simply investing in firms with an easy name. The performance of this investment strategy depends on the type of firm or the industry the agent invests in. The fluency strategy may improved by the knowledge that the fluency effect is more pronounced in certain industries and types of firms. By focusing on investing in firms with a fluent name that have, for example, a high amount of intangible assets may lead to higher excess returns. The same line of reasoning applies to firms that are active in industries that are more susceptible to the fluency effect.

A limitation of this paper is the amount of control variables that is used. By adding additional control variables, the results may become even more robust. Moreover, the sample consists only of US firms. Therefore, the study does not provide any information on countries outside the United States. The study does not control for the familiarity effect or for name changes. Moreover, this paper does not take ticker symbols into account. Many investors may make their investment decisions based on ticker symbols. This may be another limitation of the study.

For further research, I recommended to look further into the cause of the fluency effect. Even though, Multiple papers indicate that fluency has an effect on

the stock market. Furthermore, this study shows that there is variance in the bias across different industries and types of firms. However, it is not yet known what the mechanisms behind this behavioral bias are. Moreover, future papers may study what the effect of a change in fluency is on the strength of the behavioral bias.

A Fluency effect across 48 different industries

Table 17: Company name fluency and returns per industry without control variables

This table presents the estimates of regressions of excess returns on fluency. The same regression is performed on 48 subsamples that are categorized based on the Fama and French (1997) 48-industry classification. Excess returns are obtained from CRSP. *Fluency* scores are the aggregate of *length*, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0

	Agric	Food	Soda	Beer	Smoke	Toys	Fun	Books
Fluency	0.0051	0.0040	-0.0017	0.0057	0.0288	0.0034	-0.0001	0.0023
	(0.90)	(0.50)	(-0.14)	(0.48)	(3.67)	(0.33)	(-0.01)	(0.39)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	383	1,876	358	401	176	1,175	1,808	1,350
	Hshld	Clths	Hlth	MedEq	Drugs	Chems	Rubbr	Txtls
Fluency	0.0034	0.0075	0.0037	0.0048	-0.0036	0.0034	-0.0055	0.0113
	(0.64)	(1.48)	(1.22)	(1.10)	(-1.12)	(0.63)	(-1.34)	(0.81)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,109	1,469	2,954	4,218	6,483	2,310	972	643
	BldMt	Cnstr	Steel	FabPr	Mach	ElcEq	Autos	Aero
Fluency	-0.0069	0.0036	-0.0059	-0.0108	0.0044	-0.0008	0.0171	-0.0090
	(-1.43)	(0.64)	(-1.02)	(-2.78)	(1.05)	(-0.12)	(2.41)	(-0.87)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,343	1,556	1,699	407	4,172	3.413	1,681	609
	Ships	Guns	Gold	Mines	Coal	Oil	Util	Telcm
Fluency	0.0144	-0.0110	0.0033	-0.0140	0.0216	0.0038	0.0090	0.0010
	(2.50)	(-0.50)	(0.40)	(-1.41)	(0.46)	(0.85)	(5.68)	(0.34)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	242	254	426	615	237	5,006	4,047	3,532
	PerSv	BusSv	Comps	Chips	LabEq	Paper	Boxes	Trans
Fluency	0.0025	-0.0004	0.0011	0.0023	-0.0052	-0.0037	0.0259	0.0018
	(0.41)	(-0.18)	(0.25)	(0.83)	(-0.88)	(-0.43)	(7.68)	(0.44)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,432	16,063	4,308	6,923	2,543	1,332	430	2,853
	Whlsl	Rtail	Meals	Banks	Insur	RlEst	Fin	Other
Fluency	0.0034	0.0028	0.0065	0.0063	-0.0037	-0.0021	0.0014	0.0044
	(1.12)	(0.96)	(1.58)	(4.98)	(-1.12)	(-0.16)	(0.42)	(0.48)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	5,666	6,733	2,969	13,642	4,031	984	8,922	666

Table 18: Company name fluency and returns per industry with control variables

This table presents the estimates of regressions of excess returns on fluency. The same regression is performed on 48 subsamples that are categorized based on the Fama and French (1997) 48-industry classification. Excess returns are obtained from CRSP. *Fluency* scores are the aggregate of the /textitlength, *Englishness* and *dictionary* scores. All company names that consist of one, two or three words receive *length* scores of respectively 3, 2 and 1. *Englishness* is measures based on the methodology of Travers and Olivier (1978). Firms in the bottom quintile receive a score of 1 and other firms receive a score of 0. A firm gets a *dictionary* score of 1 if all words pass the Microsoft spell-check in all lower-case letters. If a word does not pass the spell-check, the name receives a *dictionary* score of 0. The decription of the control variables is included in Appendix B.

	Agric	Food	Soda	Beer	Smoke	Toys	Fun	Books
Fluency	-0.0085	0.0102	-0.0040	0.0252	0.0334	0.0010	0.0066	0.0005
	(-2.03)	(1.04)	(-0.30)	(1.68)	(3.77)	(0.06)	(0.69)	(0.06)
Log(Size)	0.0010	0.0020	-0.0075	0.0169	-0.0167	-0.0146	-0.0106	-0.0072
	(0.13)	(0.84)	(-1.38)	(3.34)	(-1.46)	(-2.80)	(-3.21)	(-1.76)
Log(B/M)	0.0073	0.0175	-0.0017	0.0328	0.0118	0.0033	0.0075	0.0023
	(0.93)	(6.08)	(-0.22)	(4.28)	(2.43)	(0.55)	(2.12)	(0.69)
Leverage	0.0023	0.0327	0.0184	0.0760	0.0972	-0.0067	0.0138	0.0271
	(0.08)	(3.08)	(0.74)	(3.15)	(1.92)	(-0.35)	(1.15)	(2.07)
Return	-0.1599	-0.1355	-0.1804	0.0412	-0.1922	-0.1070	-0.1971	-0.1363
	(-2.10)	(-4.76)	(-1.81)	(0.52)	(-2.65)	(-2.93)	(-5.43)	(-3.42)
Turnover	-0.0013	0.0021	0.0056	-0.0084	0.0264	-0.0012	-0.0022	0.0031
	(-0.50)	(0.98)	(0.63)	(-1.45)	(4.38)	(-0.35)	(-0.93)	(1.13)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	311	1,596	324	340	132	933	1,327	1,090

Table 5 Continued: Company name fluency and returns per industry with control variables

	Hshld	Clths	Hlth	MedEq	Drugs	Chems	Rubbr	Txtls
Fluency	0.0066	0.0091	0.0061	0.0073	-0.0062	-0.0022	-0.0106	0.0078
	(0.85)	(1.17)	(1.59)	(1.23)	(-1.34)	(-0.32)	(-1.81)	(0.46)
Log(Size)	-0.0066	-0.0032	-0.0102	-0.0055	-0.0074	0.0006	-0.0100	-0.0129
	(-2.28)	(-0.91)	(-4.25)	(-3.59)	(-5.19)	(0.26)	(-1.77)	(-2.54)
Log(B/M)	0.0079	0.0169	0.0098	0.0117	0.0061	0.0129	0.0089	0.0066
	(2.46)	(3.55)	(3.70)	(5.94)	(3.93)	(4.65)	(1.41)	(1.02)
Leverage	0.0214	0.0078	0.0076	0.0292	0.0281	0.0370	0.0190	-0.0423
	(1.70)	(0.75)	(0.83)	(3.67)	(4.17)	(4.67)	(1.23)	(-1.51)
Return	-0.0989	-0.0743	-0.1101	-0.1020	-0.1111	-0.1273	-0.1170	-0.0911
	(-2.33)	(-2.34)	(-5.06)	(-5.51)	(-7.99)	(-4.27)	(-2.39)	(-1.78)
Turnover	0.0001	0.0047	-0.0004	-0.0034	-0.0025	0.0023	-0.0008	-0.0004
	(0.04)	(1.89)	(-0.26)	(-2.60)	(-3.20)	(1.03)	(-0.18)	(-0.08)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,762	1,218	2,285	3,495	5,229	1,967	759	521

Table 5 Continued: Company name fluency and returns per industry with control variables

	BldMt	Cnstr	Steel	FabPr	Mach	ElcEq	Autos	Aero
Fluency	-0.0069	0.0064	-0.0094	-0.0354	0.0074	-0.0024	0.0081	-0.0114
	(-1.29)	(1.00)	(-1.91)	(-5.71)	(1.31)	(-0.32)	(1.04)	(-1.01)
Log(Size)	-0.0001	-0.0077	-0.0079	-0.0166	-0.0034	-0.0105	-0.0100	0.0017
	(-0.03)	(-2.94)	(-1.98)	(-2.71)	(-1.95)	(-4.46)	(-2.54)	(0.50)
Log(B/M)	0.0149	0.0164	0.0123	-0.0003	0.0118	0.0051	0.0072	0.0139
	(4.11)	(3.94)	(3.01)	(-0.03)	(5.07)	(1.96)	(1.96)	(2.41)
Leverage	0.0077	-0.0116	0.0175	0.0481	0.0123	0.0083	0.0094	0.0286
	(0.80)	(-0.74)	(1.29)	(1.64)	(1.67)	(0.90)	(0.68)	(1.65)
Return	-0.0531	-0.0525	-0.1582	-0.0456	-0.1166	-0.1573	-0.0427	-0.0709
	(-2.12)	(-1.40)	(-4.93)	(-0.84)	(-5.51)	(-6.95)	(-1.34)	(-1.18)
Turnover	0.0027	0.0032	0.0053	-0.0009	0.0047	-0.0046	0.0077	0.0069
	(1.21)	(2.44)	(4.04)	(-0.22)	(3.30)	(-3.86)	(3.90)	(1.57)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,910	1,270	1,458	336	3,546	2,724	1,433	509

Table 5 Continued: Company name fluency and returns per industry with control variables

	Ships	Guns	Gold	Mines	Coal	Oil	Util	Telcm
Fluency	0.0224	-0.0032	0.0062	-0.0028	0.0358	-0.0009	0.0033	-0.0004
	(3.34)	(-0.25)	(0.33)	(-0.17)	(1.70)	(-0.17)	(2.51)	(-0.08)
Log(Size)	0.0080	-0.0207	-0.0122	-0.0087	-0.0281	-0.0032	0.0097	-0.0153
	(1.60)	(-3.55)	(-2.00)	(-1.48)	(-3.44)	(-2.55)	(4.81)	(-6.31)
Log(B/M)	0.0467	0.0033	0.0148	0.0024	-0.0111	0.0145	0.0180	0.0054
	(3.86)	(0.28)	(1.57)	(0.36)	(-1.38)	(7.06)	(5.44)	(2.17)
Leverage	0.0528	-0.0226	-0.0321	0.0492	-0.0195	0.0247	0.0227	0.0355
	(1.85)	(-0.84)	(-1.46)	(2.39)	(-0.53)	(3.33)	(2.64)	(4.72)
Return	0.1146	-0.0366	-0.0798	-0.1163	-0.0585	-0.1351	-0.1539	-0.1768
	(2.43)	(-0.66)	(-0.96)	(-2.64)	(-0.66)	(-7.16)	(-4.29)	(-6.52)
Turnover	0.0028	0.0025	0.0079	0.0016	0.0102	0.0035	0.0133	-0.0007
	(1.17)	(0.39)	(1.81)	(0.47)	(2.41)	(3.43)	(6.33)	(-0.49)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	212	224	325	499	191	4,005	3,595	2,476

Table 5 Continued: Company name fluency and returns per industry with control variables

	PerSv	BusSv	Comps	Chips	LabEq	Paper	Boxes	Trans
Fluency	0.0050	-0.0009	0.0049	0.0006	-0.0110	-0.0112	0.0183	0.0010
	(0.42)	(-0.34)	(0.63)	(0.16)	(-1.55)	(-1.64)	(2.26)	(0.18)
Log(Size)	-0.0154	-0.0132	-0.0161	-0.0092	-0.0024	-0.0061	-0.0002	-0.0004
	(-3.63)	(-12.40)	(-7.85)	(-6.28)	(-1.04)	(-1.82)	(-0.03)	(-0.23)
Log(B/M)	0.0010	0.0068	0.0050	0.0072	0.0110	0.0136	0.0086	0.0083
	(0.25)	(5.73)	(2.10)	(3.71)	(3.10)	(3.50)	(1.96)	(2.96)
Leverage	-0.0286	0.0255	0.0217	0.0043	0.0303	0.0416	0.0479	0.0179
	(-2.36)	(5.70)	(2.31)	(0.67)	(2.54)	(3.11)	(2.21)	(1.79)
Return	-0.1298	-0.1604	-0.1855	-0.1587	-0.1387	-0.1111	-0.1181	-0.0992
	(-3.49)	(-14.99)	(-9.40)	(-11.29)	(-5.14)	(-2.59)	(-2.66)	(-3.38)
Turnover	-0.0010	-0.0034	-0.0012	-0.0020	0.0002	0.0108	0.0122	-0.0002
	(-0.50)	(-5.34)	(-1.05)	(-2.74)	(0.15)	(3.21)	(2.17)	(-0.16)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,091	12,642	3,538	5,938	2,147	1,164	360	2,331

Table 5 Continued: Company name fluency and returns per industry with control variables

	Whlsl	Rtail	Meals	Banks	Insur	RlEst	Fin	Other
Fluency	0.0060	0.0010	0.0095	0.0031	-0.0079	0.0025	-0.0013	0.0019
	(1.38)	(0.27)	(1.67)	(1.49)	(-2.31)	(0.12)	(-0.37)	(0.16)
Log(Size)	-0.0042	-0.0055	-0.0095	0.0039	-0.0011	-0.0076	-0.0005	0.0006
	(-2.52)	(-3.41)	(-3.80)	(4.02)	(-0.58)	(-1.93)	(-0.39)	(0.24)
Log(B/M)	0.0154	0.0183	0.0106	0.0223	0.0152	0.0132	0.0136	0.0137
	(7.89)	(10.20)	(4.20)	(13.06)	(3.10)	(2.12)	(5.86)	(2.89)
Leverage	0.0132	0.0182	0.0312	-0.0101	0.0334	0.0095	-0.0077	0.0262
	(1.59)	(2.61)	(3.72)	(-2.93)	(3.14)	(0.45)	(-1.39)	(1.12)
Return	-0.1256	-0.0382	-0.0369	0.0531	-0.1020	0.0026	-0.1261	-0.2041
	(-7.16)	(-2.23)	(-1.28)	(3.37)	(-3.93)	(0.06)	(-7.11)	(-2.37)
Turnover	-0.0018	0.0016	0.0050	-0.0018	0.0061	-0.0012	-0.0003	0.0037
	(-1.65)	(1.66)	(2.95)	(-1.68)	(4.16)	(-0.30)	(-0.24)	(0.71)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4,540	5,624	2,328	10,127	3,303	591	5,920	536

B Description of variables

Age – The number of years a security is available in CRSP

Book equity – Value of stockholders' equity plus deferred taxes and investment credit minus the redemption value of preferred stock (Kenneth French, 2018)

Book equity - Total debt divided by the sum of total debt and book equity (Hong and Kacperczyk, 2009)

Book-to-market – Book equity divided by market equity (size)

Excess returns - Stock returns minus the average annual yield of a one-month Treasury bill that year

Size – Market equity computed as stock price times shares outstanding in the prior month (Green and Jame, 2013)

Book leverage - Average annual turnover scaled by shares outstanding (Green and Jame, 2013)
(*Intangible assets* - Average annual intangible assets scaled by total book assets

C Description of the 48 industries

Agric - Agriculture
Food - Food Products
Soda - Candy & Soda
Beer - Beer & Liquor
Smoke - Tobacco Products
Toys - Recreation
Fun - Entertainment
Books - Printing & Publishing
Hshld - Consumer Goods
Clths - Apparel
Hlth - Health (services)
MedEq - Medical Equipment
Drugs - Pharmaceutical Products
Rubbr - Rubber & Plastic Products
Txtls - Textiles
BldMt - Construction Materials
Cnstr - Construction
Steel - Steel Works
FabPr - Fabricated Products
Mach - Machinery
ElcEq - Electrical Equipment
Auto - Automobiles & Trucks
Aero - Aircraft
Ships - Shipbuilding & Railroad
Guns - Defense
Gold - Precious Metals
Mines - Non-Metallic, Industrial Mining
Coal - Coal
Oil - Petroleum & Natural Gas
Util - Utilities
Telcm - Communication
PerSv - Personal Services
BusSv - Business Services
Comps - Computers
LabEq - Measuring, Control Equipment
Paper - Business Supplies
Boxes - Shipping Containers
Trans - Transportation
Whlsl - Wholesale
Rtail - Retail
Meals - Restaurants, Hotels & Motels
Banks - Banking
Insur - Insurance
RlEst - Real Estate
Fin - Trading
Other - Miscellaneous

References

Alter, A. L. and 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. and 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.

Atiase, R. K. (1985). Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research*, pages 21–36.

Bao, Y., Shao, A. T., and Rivers, D. (2008). Creating new brand names: Effects of relevance, connotation, and pronunciation. *Journal of Advertising Research*, 48(1):148–162.

Barth, M. E., Kasznik, R., and McNichols, M. F. (2001). Analyst coverage and intangible assets. *Journal of accounting research*, 39(1):1–34.

Begg, I. M., Anas, A., and Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121(4):446.

Brown, S. J., Goetzmann, W., Ibbotson, R. G., and Ross, S. A. (1992). Survivorship bias in performance studies. *The Review of Financial Studies*, 5(4):553–580.

Calantone, R. J., Cavusgil, S. T., and Zhao, Y. (2002). Learning orientation, firm innovation capability, and firm performance. *Industrial marketing management*, 31(6):515–524.

Collins, D. W., Kothari, S. P., and Rayburn, J. D. (1987). Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics*, 9(2):111–138.

Cooper, M. J., Dimitrov, O., and Rau, P. R. (2001). A rose. com by any other name. *The journal of Finance*, 56(6):2371–2388.

Cooper, M. J., Gulen, H., and Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. *The Journal of Finance*, 60(6):2825–2858.

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

Dahlquist, M. and Robertsson, G. (2001). Direct foreign ownership, institutional investors, and firm characteristics. *Journal of financial economics*, 59(3):413–440.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

Fama, E. F. and 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. and French, K. R. (1997). Industry costs of equity. *Journal of financial economics*, 43(2):153–193.

Giliberto, M. (1990). Equity real estate investment trusts and real estate returns. *Journal of Real Estate Research*, 5(2):259–263.

Grant, E. B. (1980). Market implications of differential amounts of interim information. *Journal of Accounting Research*, pages 255–268.

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

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

Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1):15–36.

Huberman, G. (2001). Familiarity breeds investment. *The Review of Financial Studies*, 14(3):659–680.

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

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The journal of finance*, 49(5):1541–1578.

Lang, M. H., Lins, K. V., and Miller, D. P. (2003). Adrs, analysts, and accuracy: Does cross listing in the united states improve a firm's information environment and increase market value? *Journal of Accounting Research*, 41(2):317–345.

Lee, C., Shleifer, A., and Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *The Journal of Finance*, 46(1):75–109.

Massa, M. and Simonov, A. (2006). Hedging, familiarity and portfolio choice. *The Review of Financial Studies*, 19(2):633–685.

Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and social psychology review*, 8(4):364–382.

Rosenberg, B., Reid, K., and Lanstein, R. (1998). market inefficiency. *Streetwise: the best of the Journal of portfolio management*, page 48.

Song, H. and Schwarz, N. (2008). If it's hard to read, it's hard to do: Processing fluency affects effort prediction and motivation. *Psychological Science*, 19(10):986–988.

Song, H. and Schwarz, N. (2009). If it's difficult to pronounce, it must be risky: Fluency, familiarity, and risk perception. *Psychological Science*, 20(2):135–138.

Travers, J. R. and Olivier, D. C. (1978). Pronounceability and statistical" englishness" as determinants of letter identification. *The American Journal of Psychology*, pages 523–538.

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