Journal of Behavioral and Experimental Finance 39 (2023) 100819

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Finance

journal homepage: www.elsevier.com/locate/jbef

Full length article

Company name fluency and stock returns☆

Maurizio Montone [a], Martijn J. van den Assem [b,c], Remco C.J. Zwinkels [b,c,*]

[a] Utrecht University, Netherlands
[b] Vrije Universiteit (VU) Amsterdam, Netherlands
[c] Tinbergen Institute, Netherlands

ARTICLE INFO

Article history:
Received 29 August 2022
Received in revised form 26 April 2023
Accepted 20 May 2023
Available online 25 May 2023

JEL classification:
G12
G14
G41

Keywords:
Fluency
Investor recognition
Company names
Stock returns

ABSTRACT

Previous research shows that stocks with fluent names trade at higher prices. In this paper, we test whether fluency simply appeals to naive investors, or actually identifies better firms. We find that companies with fluent names are more profitable, but some investors appear to neglect this information. Correspondingly, stocks with fluent names yield higher abnormal returns relative to stocks with nonfluent names. Consistent with our theoretical model, these effects are concentrated among firms with low market capitalization and high sensitivity to investor sentiment. The results lend novel support to the view that company names convey information.

1. Introduction

Stocks with fluent names trade at higher prices (Green and Jame, 2013). In this paper, we explore two competing hypotheses that can explain this phenomenon. On the one hand, the psychology literature shows that individuals judge fluent stimuli more positively than nonfluent ones (Schwarz, 2004; Oppenheimer, 2006; Song and Schwarz, 2009). Experimental and field evidence suggests that this bias may also affect financial decision-making (Alter and Oppenheimer, 2006, 2008, 2009; Shah and Oppenheimer, 2009; Silva et al., 2016; Chan et al., 2018; Schwarz et al., 2021; Hsu et al., 2022; Green and Jame, 2013) conclude that investors may indeed exhibit a naive preference, or "affect", for fluency, and bid up the prices of fluently-named stocks (henceforth: "fluent stocks").

An alternative line of reasoning is that company names convey information on the company's quality (Fombrun and Shanley, 1990; Perkins, 1995; Aaker, 1996; Tadelis, 1999). For example, the name constitutes a direct link with the company's mission statement, inner identity, or culture (see, e.g., Muzellec (2006) for a review). The value of a fluent name then lies in the fact that it makes the company's core corporate values clear to both internal and external stakeholders, which in turn increases the human intellectual capital of the company (Bart, 2001).[1] Under this scenario, fluent company names identify superior firms.

In this paper, we tell apart these two explanations by testing their respective implications for stock returns, valuations, and earnings. Building on the distinction between sophisticated and unsophisticated (or "naive") investors (Chen et al., 2002; Hirshleifer and Teoh, 2003; Hong and Stein, 2007; Hong and Sraer, 2016), the affect hypothesis implies that the latter mistakenly believe that fluent firms have more profitable projects. Their overbidding for fluent stocks leads to overpricing, and then lower returns relative to nonfluent stocks. On the other hand, the information story implies that naive investors neglect the correlation

☆ We thank Stefan Palan (the Editor), two anonymous referees, Ingolf Dittmann, David Feldman, Chris Godfrey, Clifton Green, David Hirshleifer, Russell Jame, Alberto Martín-Utrera, Paolo Porchia, Dirk Schoenmaker, Hersh Shefrin, and participants at the Israel Behavioral Finance Conference at Tel Aviv University, the Research in Behavioral Finance Conference at Vrije Universiteit (VU) Amsterdam, Leibniz Universität Hannover, the Erasmus Finance Day at Erasmus University Rotterdam, the Finance Forum at CUNEF, and the Behavioral Finance Working Group Conference at Queen Mary University of London for many helpful comments.

* Correspondence to: Vrije Universiteit Amsterdam, De Boelelaan 1105, Amsterdam, 1081 HV, Netherlands.
E-mail address: r.zwinkels@vu.nl (R.C.J. Zwinkels).

[1] Human intellectual capital is defined as the sum of human, structural, and relationship capital, which respectively represent the company's know-how, the routines and systems of the organization, and the relationships with external stakeholders (see, e.g., Bontis (1999).

https://doi.org/10.1016/j.jbef.2023.100819

between fluency and the future profitability of the firm.[2] Their underbidding for fluent stocks leads to underpricing, and higher subsequent returns.

Using fluency data on U.S. company names from Green and Jame (2013), we provide strong support for the information hypothesis. We find a positive and robust relation between company name fluency and abnormal returns. Consistent with the theory that company name fluency is informative but priced only partially, the results are concentrated among firms with a below-median market capitalization and an above-median sentiment beta. Moreover, additional analyses confirm that companies with fluent names have higher future profitability, and produce positive earnings surprises. The results lend novel support to the idea that company names convey information, a view proposed by both the management literature and practitioners in the field.[3] More generally, our findings provide a new kind of confirmation of the tendency of investors to undervalue intangibles (see, e.g., Edmans (2011)).

To derive theoretical guidance, we consider an economy from Hirshleifer and Teoh (2003). Investors are risk averse, and either sophisticated (arbitrageurs) or unsophisticated (naive investors). The distinctive feature of arbitrageurs is that they evaluate stocks correctly, whereas naive investors exhibit biases in their trading decisions. Prices can deviate from fundamental values because the finite risk tolerance of arbitrageurs limits arbitrage. In this economy, the short-term equilibrium price represents an average of the subjective evaluations of each investor type. With respect to the original setup, we introduce two elements of novelty. First, we extend the model to include multiple traded firms, each endowed with a risky project that pays a final dividend. Second, these firms are characterized by a name that can be either fluent or nonfluent.

In this modified version of the model, we explore the asset-pricing implications of the two alternative interpretations of fluency suggested by previous research. In the first setting, there is no correlation between fluency and the quality of the project, but naive investors erroneously believe that fluency identifies better stocks. In the presence of downward-sloping demand curves (Shleifer, 1986; Kaul et al., 2000), naive investor demand pushes the equilibrium prices of the two types of stocks apart, whereas in fact they should be equal. As a result, fluent stocks earn lower returns than nonfluent stocks.[4] The difference in returns increases (in absolute value) with the size of naive investor demand.

In the second setting, we explore the hypothesis that firms with fluent names have better projects. In this scenario, the projects of fluent firms yield a higher average payoff. If naive investors neglect this information, they evaluate fluent and nonfluent stocks equally. Fluent stocks then become underpriced and generate higher returns than nonfluent stocks. In this scenario, the difference in returns is positively related to the size of naive investor demand, and to the quality difference between good and bad projects.

In the empirical analysis, we take these predictions to the data. Our main analyses consider the data set of annual U.S. company name fluency scores from Green and Jame (2013), which spans the period from 1981 to 2008, and complement it with CRSP-Compustat data. The fluency score of a company name is calculated as the sum of length, ''Englishness'', and dictionary scores, and takes on integer values between zero and four. The length score attaches a higher value to shorter names, the Englishness score identifies words that are more recurrent in the English language according to the Corpus of Contemporary American English, and the dictionary score identifies words that pass the Microsoft spell-check.

The model includes two key parameters of interest. The first one is the size of naive investor demand. To identify it in the empirical tests, we acknowledge that periods of high investor sentiment are characterized by an increased presence of naive investors in the market (Baker and Wurgler, 2006, 2007; Yu and Yuan, 2011; Stambaugh et al., 2012), and that naive investors are especially active in the market for stocks with a high sentiment beta (Glushkov, 2005; Baker et al., 2012).

The second key parameter from the model is the difference in quality between fluent and non-fluent firms, either perceived (affect hypothesis) or true (information hypothesis). We empirically identify it as firm size, as quality is more heterogeneous among small companies. For example, such firms are less diversified (Frankel and Litov, 2009), and exhibit greater dispersion of managerial talent (Gabaix and Landier, 2008), which leads to greater variation in earnings.

As a preliminary test, we perform a portfolio analysis. The empirical evidence is consistent with the theoretical predictions of the information story. When Baker and Wurgler's 2007 beginning-of-period investor sentiment index is high, fluent stocks outperform nonfluent stocks by 50 basis points per month, but only among companies with below-median market capitalization. These abnormal returns are robust to the inclusion of a variety of factor-mimicking portfolios, which addresses the concern that the return differential might represent some other known asset pricing factor.[5]

One concern is that fluency may be correlated with other firm characteristics that are also known to affect stock returns. To address this point, we perform a stock-level analysis. We estimate Fama–MacBeth regressions of returns controlling for a large vector of firm characteristics from Edmans (2011), in addition to the factor-mimicking portfolios introduced above. The empirical evidence again lends support to the information story. Consistent with our theoretical model, the positive relation between fluency and stock returns is confined to firms with low market capitalization and high sensitivity to investor sentiment. Reassuringly, these return patterns are not driven by microcaps or penny stocks. We also find similar estimates in panel regressions with firm and year fixed-effects, which mitigates the concern that fluency may capture some time-invariant firm characteristics or market-wide time trends. Finally, we find similar results for each of the three individual components of the fluency index.

Next, we turn to the model predictions on valuations. Under the information hypothesis, a larger participation of naive investors should bring the prices of fluent and nonfluent stocks closer to each other, especially among firms that are more heterogeneous in quality. To test this conjecture, we estimate valuation regressions from Hong and Kacperczyk (2009), and augment them with our variables of interest. Consistent with the information hypothesis, we find that the price differential between fluent and nonfluent stocks decreases with sentiment beta. The effect is limited to small stocks, as in our previous results.

2 The main reason why investors may neglect a relevant variable in their evaluations is that they lack information on its value (the so-called ''lack-of-information'' hypothesis). See Edmans (2011) for an excellent review of this literature.

3 For example, there are plenty of consultants that help startups pick a name, and the characteristics they favor are strikingly similar to the measure of company name fluency we consider (see, e.g., Harroch, ''12 Tips For Naming Your Startup Business'', *Forbes*, 2016).

4 The inverse relation between affect and stock returns is well-known in the asset pricing literature (see, e.g., Statman et al. (2008).

5 The factors we consider are the market, size, book-to-market, investment, and profitability factors from Fama and French (2015), the momentum factor from Carhart (1997), the liquidity factor from Pástor and Stambaugh (2003), the size, management, and performance factors from Stambaugh and Yuan (2017), and the book-to-market and profitability factors from Novy-Marx (2013).

Harvey (2017) makes the compelling point that research in asset pricing can produce spurious results if the underlying economic mechanism is unclear. To address this concern, we test the two underpinnings of the information hypothesis. First, we confirm that there is a strong positive relation between company name fluency and future operating performance. This result is important because under the alternative affect hypothesis there is no correlation between fluency and profitability. Second, we provide evidence that the mispricing of fluency is related to expectation errors. Specifically, we show that fluent firms systematically surprise analysts with unexpected positive earnings. Both sets of findings are again more pronounced among small companies, and suggest that fluency indeed conveys information on the quality of the firm.

The above analyses use the same fluency scores and considered the same sample period as in Green and Jame (2013). In our final group of tests, we analyze whether our main findings also hold over a longer sample period that also includes more recent times. To this end, we construct an updated version of the fluency index that extends until 2021 using the most recent versions of the Corpus and the Microsoft spell-check. We find again that fluent stocks with high sentiment betas exhibit higher abnormal returns than their non-fluent counterparts, but the results are confined to the length component of the fluency index. These findings seems to reflect two drawbacks of the Englishness and dictionary measures derived from the two new data sources. First, they may not fully apply to the earlier part of our sample period because language has evolved. Second, they do not include recent language changes driven by social media (e.g., alternative spellings, neologisms). The length dimension, on the other hand, remains unaffected by either of these issues.

Our findings speak to a growing literature on fluency and stock returns. Alter and Oppenheimer (2006) find a positive association between fluent ticker codes and IPO returns, but do not use standard methods and controls from the asset pricing literature. In a related study, Head et al. (2009) consider a limited sample of U.S. stocks and find that fluent ticker codes are also associated with higher long-run returns.[6] Jin et al. (2021) analyze China's A-share market, and find that stocks with three-character names yield higher abnormal returns than stocks with four-character names. In this paper, we complement these studies by analyzing the universe of U.S. stocks. Consistent with their findings, we observe a positive relation between company name fluency and stock returns.

We also contribute to the literature on fluency and stock valuations. Green and Jame (2013) find a strong positive relation between measures of company valuations, such as market-to-book and Tobin's q, and company name fluency. They also re-estimate this relation in subsamples of firms of different size, and find that the effect is more pronounced for firms with low market capitalization. In our paper, we extend their analysis in two ways. First, we provide formal statistical evidence that the effect of fluency on valuations is indeed stronger for small firms. Second, we show that the effect is concentrated around firms with a high sentiment beta. Both findings lend support to the theoretical mechanism we propose.

Green and Jame (2013) further show that fluent stocks also exhibit higher breadth of ownership and liquidity, and propose the affect hypothesis as the mechanism that underlies their results. In this paper, we propose the information hypothesis as an alternative explanation. To tease out the two explanations, we develop a formal theoretical model that nests both of them. Central in distinguishing between the two hypotheses is their different impact on expected returns.[7] We corroborate the (untabulated) observation from Green and Jame (2013) that there is no unconditional relationship between fluency and expected returns. Guided by our theoretical model, however, we extend the analysis to show that there is a conditional positive relationship between fluency and expected returns that holds for firms with a high sentiment beta and smaller size. Therefore, the findings ultimately lend support to the information hypothesis.

Our paper also speaks to an established literature that shows that investors tend to underprice intangibles, such as employee satisfaction (Edmans, 2011), R&D (Lev and Sougiannis, 1996; Chan et al., 2001), innovative efficiency (Hirshleifer et al., 2018), advertising (Chan et al., 2001), patent citations (Deng et al., 1999), and software development costs (Aboody and Lev, 1998). Stocks that score high on these characteristics earn higher long-run returns. The main explanation for these results is that investors neglect a relevant variable if its value is not clear, or if it is costly to obtain or construct (see Edmans (2011)). Our results suggest that a similar mechanism applies to company name fluency.

Overall, our results show that fluency conveys information on the quality of the company. While the information embedded in company name fluency may overlap with measures of attention, brand recognition, media coverage, or other intangible assets, the crucial advantage of using fluency is that it is less likely to be endogenously affected by firm performance or investment. As a result, fluency can be considered as a sufficient and reasonably exogenous measure of firm quality. Finally, we acknowledge that the affect and the information channel are not necessarily mutually exclusive, and might coexist. While our results cannot definitively exclude the presence of the affect channel, they suggest that the information channel dominates.

The paper proceeds as follows. Section 2 introduces the theoretical framework. Section 3 describes the data. Section 4 presents the empirical results. Section 5 concludes.

2. Model

We consider an economy from Hirshleifer and Teoh (2003), where investors exhibit mean–variance preferences and can be either sophisticated (arbitrageurs) or unsophisticated (naive investors). We modify the original setting in two ways. First, we extend the model to n publicly traded firms, each endowed with a risky project that pays a final dividend. Second, each of these firms is characterized by a name that is either fluent or nonfluent.

The distinctive feature of arbitrageurs is that they evaluate stocks correctly. Naive investors, instead, are prone to expectation errors. We define the probability that an investor fails to identify and process some aspect of the economic environment correctly as $f(c)$, with $f'(c) < 0$, where c represents the resources expended on attending to relevant information. Function f, then, also represents the proportion of naive investors in the economy (see Hirshleifer and Teoh (2003)). We take f as exogenously given.

The economy has three dates. At date 0 investors form expectations. At date 1, public information arrives about firm value or its components. At date 2 the terminal payoff is realized and the firm is liquidated. There is no private information among investors, so there is nothing to learn from the market price in this economy. Naive investors, however, are not aware that they are not processing information fully, so they mistakenly believe that they too have nothing to learn. Therefore, they do not update their beliefs based upon the market price.

6 It is important to note that ticker codes are rather brief in comparison with company names, and can even generate confusion among investors (see, e.g., Rashes (2001)). However, ticker codes are typically considered to be fluent, or "clever", if they exhibit a clear relation to the company's business (see, e.g., Head et al. (2009)).

7 The model is silent on breadth of ownership and stock liquidity.

We assume that each investor is endowed with an initial wealth of W_0 and x_{0i} units of each risky security i. At date 1, investors can buy or sell securities in exchange for cash, defined as claims to terminal consumption C, at price S_{1i}. We denote the position in security i thus attained as x_i, and the terminal payoff of the security as S_{2i}. Then an individual's consumption is $C = W_0 + \sum_{i=1}^{n} x_{0i}S_{1i} + \sum_{i=1}^{n} x_i(S_{2i} - S_{1i})$.

An investor of type ϕ solves (see Appendix for details):

$$\max_{\{x_i^{\phi}\}_{i=1}^{n}} E_1^{\phi}\left(\sum_{i=1}^{n} x_i^{\phi}(S_{2i} - S_{1i})\right) - \frac{\gamma}{2} var_1^{\phi}\left(\sum_{i=1}^{n} x_i^{\phi} S_{2i}\right), \tag{1}$$

where the index ϕ indicates arbitrageurs ($\phi = A$) or naive investors ($\phi = N$), and γ is the coefficient of absolute risk aversion. The market clearing condition for security i is $fx_i^N + (1-f)x_i^A = x_{0i}$, where x_{0i} is the security's net supply. The equilibrium price of security i is then:

$$S_{1i} = fE_1^N(S_{2i}) + (1 - f)E_1^A(S_{2i}), \tag{2}$$

which represents a weighted average of the beliefs of the two types of investors. The intuition behind this result is that arbitrageurs exhibit a finite level of risk-tolerance, which represents a limit to arbitrage (see Hirshleifer and Teoh (2003)). Therefore, prices do not immediately converge to the fundamental value.

We consider two scenarios. In the first one, there is no actual correlation between fluency and the quality of the firm, and all firms pay out a unit dividend. However, naive investors erroneously perceive fluent stocks to be better companies. Their evaluations of the final payment are $b_H > 1$ and $b_L \in (0, 1)$ for stocks with high and low fluency, respectively. Conversely, sophisticated investors correctly estimate the expected final payment as equal to one. This implies (see Appendix):

Proposition 1. *Under the affect hypothesis, the price of fluent stocks is higher than that of nonfluent stocks. The price differential is proportional to the size of naive investor demand and the fluency bias, and becomes zero in the absence of naive investors:*

$$S_1^H - S_1^L = f(b_H - b_L). \tag{3}$$

The equilibrium prices of fluent and nonfluent stocks should be equal, but naive investor demand drives the two prices apart. This implies:

Proposition 2. *Under the affect hypothesis, returns on fluent stocks are lower than those on nonfluent stocks. The return differential is proportional to the size of naive investor demand and the fluency bias, and becomes zero in the absence of naive investors:*

$$E(\tilde{r}_2^H) - E(\tilde{r}_2^L) = -f(b_H - b_L). \tag{4}$$

In the alternative setting, the level of fluency of the company name actually conveys information on the quality of the project. Fluent firms make an expected final payment of $\mu\lambda$, with $\lambda > 0$ and $\mu > 1$, whereas nonfluent firms make an expected final payment of λ. Arbitrageurs acknowledge the quality difference between fluent and nonfluent firms. Naive investors neglect this information, and mistakenly evaluate the payoff to be λ for all firms. We derive (see Appendix):

Proposition 3. *Under the information hypothesis, the price of fluent stocks is higher than that of nonfluent stocks. The price differential decreases with the size of naive investor demand, and increases with the systematic quality difference between fluent and nonfluent firms:*

$$S_1^H - S_1^L = \lambda(1 - f)(\mu - 1) \tag{5}$$

The equilibrium price of fluent stocks is higher than that of nonfluent stocks, but naive investor demand brings the two equilibrium prices closer to each other than they should be. This implies:

Proposition 4. *Under the information hypothesis, fluent stocks yield higher returns than nonfluent stocks. The return differential increases with the size of naive investor demand and with the systematic quality difference between fluent and nonfluent firms, and becomes zero in the absence of naive investors:*

$$E(\tilde{r}_2^H) - E(\tilde{r}_2^L) = \lambda f(\mu - 1). \tag{6}$$

A key underlying assumption for our analysis is that fluency affects naive investors. In our model, both potential channels – affect and information – rely on their actions. The exact form of naiveté, however, is different across the two channels. Under the affect hypothesis, naive investors are drawn towards fluent firms, thereby making them overvalued. Under the information hypothesis, naive investors neglect the information captured by fluency, thus making fluent firms undervalued. The group of naive investors from our model, then, is heterogeneous in the sense that one subtype of naive investors is drawn to fluent stocks, whereas the other subtype neglects the information embedded in fluency. Which of the two groups dominates is ultimately an empirical question, depending on their net effect on stocks returns. While the affect channel implies lower returns for fluent stocks, the information channel implies higher returns.[8]

In the empirical analysis that follows, we take the model's predictions to the data.

3. Data

We use the data set of U.S. company name fluency scores from Green and Jame (2013), which covers the sample period from 1981 to 2008 for all common stocks in CRSP. The overall fluency score of a firm is the sum of three variables. The first (and main) component is length, defined as the number of words in a company name, ignoring articles, conjunctions, the state of incorporation, hyphens, and expressions that are an official but often omitted part of the legal name (such as Corp., Inc., Ltd., LLC, and FSB). The length score respectively takes on a value of two, one, or zero, depending on whether the company name is made up of one word, two words, or more than two words.[9]

The other two components are Englishness and dictionary scores. The Englishness variable is assessed using the linguistic algorithm by Travers and Olivier (1978), and regressed on name length due to their high correlation. The residuals of this regression are then ranked in quintiles, and a company name takes on an Englishness score of zero if it lies in the bottom quintile of (residual) Englishness, and one otherwise. Finally, the dictionary variable takes on value one if all words in the name pass the Microsoft spell-check and zero otherwise. Overall, the index takes on integer values between zero and four. Scores are recorded on

[8] Note that the category of naive investors includes individuals and institutions alike, as both investor types are prone to biases (see, e.g., Barber and Odean (2002), Hong and Kostovetsky (2012)). In fact, a growing body of research shows that the dichotomy between highly sophisticated traders and less sophisticated ones from theoretical models seems to map into hedge funds and mutual funds, respectively (see, e.g., Chen et al. (2002); Hong and Sraer, 2013, 2016). For example, mutual funds are the driving force behind Baker and Wurgler's 2006 investor sentiment index (DeVault et al., 2019), which captures changes in stock demand not explained by economic fundamentals.

[9] In Green and Jame (2013), these scores are respectively set to three, two, and one. Although it is a purely cosmetic change, we set the minimum fluency value to zero rather than one. The reason is to ensure that a long company name earns no points in the same way as a name with a low Englishness or dictionary score.

Table 1
Summary statistics: Fluency measures.

Panel A

Variable	Mean	St. Dev.	Min	Max
Fluency	2.14	0.88	0	4
Length	1.01	0.71	0	2
Dictionary	0.33	0.47	0	1
Englishness	0.81	0.40	0	1

Panel B

Variable	Scores	Frequency	Percent	Cumulative
Fluency	0	4,202	3.29	3.29
	1	24,909	19.50	22.79
	2	50,719	39.71	62.50
	3	44,161	34.58	97.08
	4	3,729	2.92	100.00
Length	0	31,272	24.48	24.48
	1	63,801	49.95	74.44
	2	32,647	25.56	100.00
Dictionary	0	85,985	67.32	67.32
	1	41,735	32.68	100.00
Englishness	0	24,804	19.42	19.42
	1	102,916	80.58	100.00

Panel C

Industry	Mean	St. Dev.	Min	Max
Agriculture	2.11	0.98	0	4
Construction	2.20	0.86	0	4
Manufacturing	2.28	0.84	0	4
Communications	2.00	0.90	0	4
Trade	2.18	0.87	0	4
Finance	1.83	0.89	0	4
Services	2.23	0.85	0	4
Public Administration	2.24	0.78	1	3

Summary statistics for the overall fluency score and its components (Panel A), distribution across scores (Panel B), and fluency scores across industries (Panel C) for the U.S. company name fluency scores from Green and Jame (2013). The overall fluency score is split into its three components: Length, Dictionary, and Englishness. Scores are recorded on December 31 of each calendar year. The sample period is from 1981 through 2008, for an overall number of 127,720 firm-year observations.

Table 2
Summary statistics: Firm-level characteristics and portfolio returns.

Panel A. Firm characteristics

Variable	Mean	St. Dev.	p25	Median	p75
Total Assets ($ millions)	7,673	50,411	214	757	2,880
Net Sales ($ millions)	3,302	11,907	150	562	2,125
EBITDA ($ millions)	610	2,585	20	75	310
EBIT ($ millions)	462	2,160	14	54	224
Dividends ($ millions)	0.74	0.88	0.24	0.52	1.00
Closing Price ($)	29.28	30.54	15.25	24.25	36.50
Price-to-Dividend	84.74	209.64	27.42	44.82	82.21
Tobin's q	1.54	0.94	1.05	1.23	1.68
Market-to-Book	2.46	8.62	1.22	1.72	2.58
EBITDA/Total Assets	0.13	0.09	0.07	0.13	0.18
EBIT/Total Assets	0.10	0.08	0.04	0.09	0.13
R&D-to-Sales	0.03	0.04	0.00	0.02	0.04

Panel B. Firm characteristics by fluency score

Variable	Score 0	Score 1	Score 2	Score 3	Score 4
Total Assets ($ millions)	6,520	8,736	7,774	8,417	5,107
Net Sales ($ millions)	2,971	2,943	3,193	3,976	4,090
EBITDA ($ millions)	477	604	620	697	658
EBIT ($ millions)	355	493	473	504	453
Dividends ($ millions)	0.85	0.75	0.72	0.73	0.74
Closing Price ($)	30.54	28.65	29.87	29.39	37.43
Price-to-Dividend	74.18	80.15	80.37	92.20	79.68
Tobin's q	1.44	1.47	1.54	1.64	1.55
Market-to-Book	2.32	2.27	2.44	2.66	2.99
EBITDA/Total Assets	0.11	0.11	0.13	0.14	0.15
EBIT/Total Assets	0.08	0.09	0.10	0.11	0.11
R&D-to-Sales	0.02	0.03	0.03	0.03	0.02

Panel C. Portfolio returns

Variable	Mean	St. Dev.	p25	Median	p75
Score 0	0.0098	0.0477	−0.0176	0.0124	0.0385
Score 1	0.0112	0.0421	−0.0140	0.0113	0.0375
Score 2	0.0124	0.0444	−0.0158	0.0154	0.0391
Score 3	0.0125	0.0476	−0.0146	0.0144	0.0422
Score 4	0.0141	0.0558	−0.0197	0.0158	0.0425
Long-short 4–0	0.0043	0.0415	−0.0155	0.0023	0.0197
Scores 0,1	0.0110	0.0422	−0.0161	0.0119	0.0370
Scores 3,4	0.0127	0.0478	−0.0159	0.0143	0.0417
Long-short 3,4-0,1	0.0016	0.0189	−0.0095	0.0018	0.0121

Summary statistics for the firm-level characteristics (Panels A and B) and portfolio returns (Panel C) in our sample. The list of firm characteristics includes total assets; net sales, defined as the amount of billings to customers for regular sales completed during the period reduced by cash discounts, trade discounts, and returned sales and allowances for which credit is given to customers; EBITDA; EBIT; common dividends; the closing stock price; the price-to-dividend ratio; Tobin's q, defined as enterprise value (debt plus market value of equity) divided by book value (debt plus book value of equity); the market-to-book ratio, defined as market capitalization divided by the book value of equity; return on assets, defined as either EBITDA or EBIT divided by the book value of assets; and total R&D expenses divided by total sales. Panel A presents the mean, standard deviation, and 25th, 50th, and 75th percentiles for the full sample, while Panel B reports the mean value of each variable for firms with different beginning-of-year company name fluency scores. In Panel C, we report the summary statistics for value-weighted returns of portfolios with a long position in stocks with fluency scores of 0, 1, 2, 3, and 4, respectively, a portfolio with a long position in stocks with a fluency score of 4 and a short position in stocks with a fluency score of 0, a portfolio with a long position in stocks with fluency scores of 0 and 1 combined, or 3 and 4 combined, and a portfolio with a long position in stocks with a fluency score of 3 or 4 and a short position in stocks with a fluency score of 0 or 1. Accounting variables refer to the end of the fiscal year, whereas stock prices and fluency scores are measured on December 31 of each calendar year. Firm-level data is from Compustat, stock-level data is from CRSP, and fluency data is from Green and Jame (2013). The sample period is from 1981 to 2007, for an overall number of 127,720 firm-year observations.

December 31 of each calendar year from 1981 to 2008. The total number of firm-year observations is 127,720.[10]

Table 1, Panel A, presents the summary statistics for the fluency measure and its components. The mean score is 1.01 for length, 0.33 for dictionary, 0.81 for Englishness, and 2.14 for the overall fluency index. In unreported analyses we find that this value is quite stable over time, ranging between 2.10 and 2.19. Panel B shows that the distribution of fluency scores across stocks is roughly symmetric. The categories with the fewest observations are those with extreme fluency scores, i.e., zero (4,202) and four (3,729). The category with the most observations is score two (50,719), followed by scores three (44,161) and one (24,909). In Panel C, we find no major differences in average fluency scores at the industry level.

In Table 2, Panel A, we present the summary statistics for the firm-level variables. Accounting variables are from Compustat, and refer to the end of the fiscal year. Market prices are measured on December 31 of each calendar year. Across all firm-years, the average company exhibits total assets of $7.7 billion, net sales of $3.3 billion, EBITDA of $0.61 billion, and EBIT of $0.46 billion. The average Tobin's q and market-to-book ratio are 1.54 and 2.46, respectively. The average return on assets, defined as EBITDA (EBIT) divided by the book value of assets, is 13% (10%).

In Panel B, we break down these mean values into different beginning-of-year fluency scores, whereas in Panel C we report the summary statistics for portfolio returns on stocks with different fluency scores. Overall, we find that firms with higher fluency scores tend to exhibit higher Tobin's q, market-to-book, and return on assets, and yield higher stock returns. In the empirical analysis that follows, we shed further light on these patterns.

10 In the empirical tests that follow, the number of observations decreases because not all variables are available for all firms and time periods.

4. Empirical results

We present our empirical findings as follows. First, we estimate abnormal returns using factor models. Second, we estimate Fama–MacBeth regressions of returns using a large set of firm characteristics and factor loadings. Third, we estimate panel regressions of returns with firm and year fixed-effects. Fourth, we repeat the analyses of returns by separately considering the three individual components of the fluency index. Fifth, we test the model predictions on valuations. Sixth, we further explore the economic mechanism that underlies the results through an analysis of profitability and earnings surprises. Finally, we analyze the relation between fluency and stock returns in an updated sample that extends until 2021.

4.1. Time series regressions

As a preliminary test, we tease out the affect and the information story through a portfolio analysis of abnormal returns:

$$R_t = \alpha + \beta' Z_t + \epsilon_t, \tag{7}$$

where the dependent variable is the value-weighted return of a portfolio with a long position in fluent stocks (scores three or four) and a short position in nonfluent stocks (scores zero or one), with annual rebalancing, α captures abnormal returns, and Z_t is a vector of factor-mimicking portfolios, including the market, size, book-to-market, investment, and profitability factors from Fama and French (2015), the momentum factor from Carhart (1997), the liquidity factor from Pástor and Stambaugh (2003), the size, management, and performance factors from Stambaugh and Yuan (2017), and the book-to-market and profitability factors from Novy-Marx (2013).[11]

The two key parameters of interest from the model are the size of naive investor demand and the difference in quality between fluent and non-fluent firms. In the empirical analysis, we identify them as follows. With regard to the former, we build on the insight that unsophisticated investors trade when they are bullish, and leave the market when they are bearish (see, e.g., Grinblatt and Keloharju (2001); Lamont and Thaler (2003); Amromin and Sharpe (2009); Antoniou et al. (2016)). As a result, mispricing arises in times of high sentiment and disappears in times of low sentiment (Yu and Yuan, 2011; Stambaugh et al., 2012; Antoniou, Doukas, and Subrahmanyam, 2016).

In light of these considerations, we expect abnormal returns on our long–short portfolio to be larger (in absolute value) after high-sentiment periods. The intuition is as follows. The short-term equilibrium price is a weighted average of the subjective evaluations of sophisticated and unsophisticated investors (see Eq. (2)). When sentiment is high, the fraction of unsophisticated investors in the market increases, and therefore so does the impact of their demand on the price of fluent stocks. If their evaluation of fluent stocks is relatively high (low) with respect to that of arbitrageurs, then high-sentiment periods should be followed by low (high) returns on our long–short portfolio.

As for the second parameter of interest, we argue that quality is more heterogeneous among small companies. For example, such firms are less diversified (Frankel and Litov, 2009), and exhibit greater dispersion of managerial talent (Gabaix and Landier, 2008), which leads to greater variation in earnings. In the information hypothesis, this implies that the positive relation between fluency and abnormal returns should be particularly strong among small firms. Conversely, differences in firm quality have no impact on abnormal returns in the affect story, because fluency does not identify better firms.

Following these insights, we split the sample into months in which the beginning-of-period investor sentiment index from Baker and Wurgler (2007), expressed in changes and orthogonalized to business cycle indicators, is high and low, respectively, and test whether abnormal returns are significantly different across these periods.[12] The information (affect) hypothesis implies $\alpha > 0$ ($\alpha < 0$) in periods of high sentiment, and $\alpha = 0$ when sentiment is low. In a similar vein, we also consider subsamples of stocks with below- and above-median market capitalization, respectively. In the information story, the underpricing of fluent stocks should be stronger among small firms.

The results are in Table 3. Consistent with the information hypothesis, we find that the effect of fluency on stock returns is positive and significant, and only present among small stocks. Specifically, the small-stock long–short portfolio earns monthly abnormal returns of 0.28% (t-stat 2.21), whereas the alpha is not significant for the other portfolios. Furthermore, the results are entirely confined to times of high beginning-of-period sentiment. Following high sentiment, the abnormal returns on the small-stock long–short portfolio are equal to 0.64% (t-stat 4.87), whereas they are close to zero in both magnitude and significance when sentiment is low (−0.09%, t-stat −0.46).

We also re-estimate the test equation by including a dummy variable that takes on value one if beginning-of-month sentiment is high, and zero otherwise. The results are similar. Among small stocks, the coefficient of the dummy variable is equal to 0.46% (t-stat 3.54), and completely absorbs the explanatory power of the regression constant. On the other hand, the coefficients are not significant when considering all stocks or the subsample of large stocks. Therefore, these additional results provide further support to the findings from the sample breakdown, indicating that fluent stocks outperform nonfluent stocks only following periods of high sentiment, and only among stocks with below-median market capitalization.

These findings support the information hypothesis, and their robustness to known factors suggests that fluency constitutes a novel effect.[13]

4.2. Fama–MacBeth regressions

One concern is that fluency may be correlated with other firm characteristics that are also known to affect stock returns. To address this point, we estimate the following Fama–MacBeth regressions from Edmans (2011):

$$R_{i,t} = \beta_1 F_{i,y-1} + \beta_2 S_{i,t-1} + \beta_3 S_{i,t-1} \times F_{i,y-1} + \gamma' Z_{i,t} + \epsilon_{i,t}, \tag{8}$$

where $R_{i,t}$ is the excess return on stock i in month t; $F_{i,y-1}$ is the beginning-of-year company name fluency score; $S_{i,t-1}$ is beginning-of-period sentiment beta of stock i, and $Z_{i,t}$ is a vector of firm characteristics.[14]

[11] We aggregate the top two and bottom two fluency scores, respectively, to make sure we have enough stocks in each portfolio.

[12] Baker and Wurgler's 2007 investor sentiment measure is based on a number of sentiment proxies suggested in previous literature, including the closed-end fund discount, NYSE share turnover, the number and average first-day returns of IPOs, the equity share in new issues, and the dividend premium.

[13] For example, the inclusion of Pástor and Stambaugh's 2003 liquidity factor also rules out the concern that the difference in returns might represent a liquidity premium (Green and Jame, 2013; Anderson and Larkin, 2019).

[14] Edmans (2011) uses this empirical model, borrowed from Brennan et al. (1998), to analyze the relationship between employee satisfaction and long-run stock returns. The vector of characteristics includes firm size, defined as the log of market capitalization at the end of month $t-2$; the log of the book-to-market ratio, calculated each July and held constant through the following June; the ratio of dividends in the previous fiscal year to market value at calendar year-end, calculated each July and held constant through the following June; cumulative returns over months $t-3$ through $t-2$, months $t-6$ through $t-4$, and months $t-12$ through $t-7$; the log of the dollar volume of trading in the stock in month $t-2$; the log of the stock price at the end of month $t-2$.

Table 3
Abnormal returns on portfolios formed on dual fluency scores.

Panel A. All stocks

Dep. Var.: Long-short 3,4-0,1	(1) Full	(2) High	(3) Low	(4) Full
High sentiment (−1)				0.0019 (1.28)
Alpha	0.0008 (0.81)	0.0014 (0.88)	0.0001 (0.12)	−0.0003 (−0.28)
Controls	Y	Y	Y	Y
Observations	336	165	171	336
Adj. R-squared	0.4299	0.3470	0.5327	0.4307

Panel B. Small stocks

Dep. Var.: Long-short 3,4-0,1	(1) Full	(2) High	(3) Low	(4) Full
High sentiment (−1)				0.0046*** (3.54)
Alpha	0.0028** (2.21)	0.0064*** (4.87)	−0.0009 (−0.46)	0.0004 (0.27)
Controls	Y	Y	Y	Y
Observations	336	165	171	336
Adj. R-squared	0.4332	0.3900	0.5038	0.4467

Panel C. Large stocks

Dep. Var.: Long-short 3,4-0,1	(1) Full	(2) High	(3) Low	(4) Full
High sentiment (−1)				0.0019 (1.23)
Alpha	0.0008 (0.80)	0.0014 (0.84)	0.0002 (0.19)	−0.0002 (−0.25)
Controls	Y	Y	Y	Y
Observations	336	165	171	336
Adj. R-squared	0.4229	0.3411	0.5252	0.4235

OLS regressions of value-weighted excess returns on a portfolio with a long position in fluent stocks and a short position in nonfluent stocks. The regressors are the market, size, book-to-market, investment, and profitability factors from Fama and French (2015), the momentum factor from Carhart (1997), the liquidity factor from Pástor and Stambaugh (2003), the size, management, and performance factors from Stambaugh and Yuan (2017), and the book-to-market and profitability factors from Novy-Marx (2013). The fluency scores are from Green and Jame (2013), refer to company names, and take on integer values between 0 (least fluent) to 4 (most fluent). The long leg of the portfolio includes stocks with a fluency score of 3 or 4, while the short leg includes stocks with a fluency score of 0 or 1. The portfolios are formed on December 31 of the previous calendar year. We consider the all stocks in Panel A, stocks with below-median market capitalization in Panel B, and stocks with above-median market capitalization in Panel C. In columns (2) and (3), we split the sample into months in which the beginning-of-period investor sentiment index from Baker and Wurgler (2007), orthogonalized to business cycle indicators, is positive and negative, respectively. In column (4), we include a dummy variable that takes on value one if beginning-of-period sentiment is positive, and zero otherwise. Stock data is from CRSP. Observations are monthly, the sample period is from January 1982 to December 2008, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Sentiment betas measure how sensitive a stock is to market-wide sentiment, and represent the OLS coefficient of rolling five-year regressions (starting five years prior to the sample period) of monthly excess stock returns on Baker and Wurgler's 2007 investor sentiment index, expressed in changes, and orthogonalized to business cycle indicators, controlling for excess returns on the market portfolio.[15] Therefore, naive investors account for a large fraction of trading in the market for stocks with a high sentiment beta. All sentiment betas are winsorized at the 1% and 99% of the distribution, and standardized by subtracting their mean and dividing by their standard deviation to ease the interpretation of the estimates. The information (affect) story implies $\beta_1 = 0$ and $\beta_3 > 0$ ($\beta_3 < 0$).

The results are in Table 4. In column (1), we analyze the unconditional effect of fluency on stock returns by leaving sentiment out of the test equation. We find that the coefficient of the fluency index as a standalone variable is not significant. Green and Jame (2013) report a similar result in untabulated tests, and conclude that there is no effect of fluency on stock returns.[16]

Previous asset pricing research identifies a number of firm characteristics that do not have any unconditional predictive power, but actually display predictive ability after conditioning on sentiment (Baker and Wurgler, 2006, 2007; Baker et al., 2012). In light of this, we test whether a similar mechanism applies to company name fluency. The results are in column (2). Consistent with the information story, we find that sentiment beta significantly increases the effect of fluency on stock returns. For stocks

15 See Glushkov (2005). Hong and Kacperczyk (2009) propose a similar procedure to include a pure time series variable (the market factor) in Fama–MacBeth regressions of this kind.

16 However, their analysis of returns is only exploratory (see Green and Jame (2013, page 823, Section 5.3).

Table 4
Fama–MacBeth regressions of returns, fluency scores, and investor sentiment.

Dep. Var.: Ri-Rf	(1)	(2)	(3)	(4)
Fluency	−0.0002	−0.0002	−0.0004	−0.0004
	(−1.11)	(−0.74)	(−1.34)	(−1.46)
Sentiment Beta		−0.0015	−0.0014	−0.0010
		(−1.35)	(−1.16)	(−0.81)
Fluency × Sent. Beta		0.0004**	−0.0002	−0.0001
		(2.10)	(−0.64)	(−0.32)
Fluency × Sent. Beta × High Beta			0.0013***	0.0011***
			(3.69)	(4.76)
High Beta			0.0000	0.0006
			(0.02)	(0.60)
Fluency × High Beta			−0.0003	−0.0004
			(−0.67)	(−0.89)
Book-to-Market (−1)	0.0015**	0.0013**	0.0013**	−0.0012**
	(2.57)	(2.25)	(2.42)	(−2.27)
Dividend Yield (−1)	0.0008*	0.0008*	0.0008*	−0.0002
	(1.83)	(1.83)	(1.85)	(−0.76)
CumRet (−2,-3)	0.0018	0.0024	0.0022	0.0036
	(0.47)	(0.65)	(0.60)	(0.51)
CumRet (−4,-6)	0.0055	0.0072**	0.0069**	0.0120**
	(1.59)	(2.33)	(2.22)	(2.41)
CumRet (−7,-12)	0.0102***	0.0110***	0.0109***	0.0195***
	(3.49)	(4.20)	(4.18)	(6.44)
Size (−2)	−0.0003	−0.0003	−0.0002	−0.0011***
	(−0.48)	(−0.54)	(−0.42)	(−3.10)
Price (−2)	−0.0018***	−0.0018***	−0.0017**	0.0017***
	(−2.70)	(−2.71)	(−2.46)	(4.03)
Volume (−2)	−0.0000	−0.0001	−0.0002	0.0001
	(−0.04)	(−0.20)	(−0.40)	(0.24)
Factor loadings	N	N	N	Y
Observations	355,663	355,663	355,663	354,555
R-squared	0.0662	0.0944	0.0984	0.5074

Fama–MacBeth regressions of U.S. monthly excess stock returns on company name fluency scores, investor sentiment beta, a dummy variable that takes on value one if the investor sentiment beta is positive, interaction terms between fluency, sentiment beta, and the dummy, a set of firm characteristics from Edmans (2011), and factor loadings for the market, size, book-to-market, investment, and profitability factors from Fama and French (2015), the momentum factor from Carhart (1997), the liquidity factor from Pástor and Stambaugh (2003), the size, management, and performance factors from Stambaugh and Yuan (2017), and the book-to-market and profitability factors from Novy-Marx (2013), estimated over a rolling three-year window as in Hong and Kacperczyk (2009). The fluency scores are from Green and Jame (2013), measured on December 31 of each calendar year, and take on integer values between 0 (least fluent) and 4 (most fluent). We calculate sentiment betas as the winsorized (at the 1% tails) OLS coefficient of rolling five-year regressions of monthly excess stock returns on Baker and Wurgler's 2007 investor sentiment index, orthogonalized to business cycle indicators, controlling for excess returns on the stock market portfolio. To ease the interpretation of the estimates, all sentiment betas are standardized by subtracting their mean and dividing by their standard deviation. Stock data is from CRSP, accounting data is from Compustat, and the investor sentiment index is from Jeffrey Wurgler's website. The sample period is from January 1982 to December 2008, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

that exhibit a one-standard-deviation sentiment beta, the effect of a one-point increment in the fluency score on excess stock returns increases by 0.04%.

Next, we test the prediction that the mispricing of firms with fluent company names is confined to stocks with a large proportion of naive investors. To this end, we introduce a dummy variable that takes on value one for stocks with a positive sentiment beta, and zero otherwise, and interact it with our variables of interest. The intuition is that stocks that exhibit a positive and large sentiment beta are primarily traded by less sophisticated investors (Glushkov, 2005).

The results are in column (3). Consistent with the conjecture, we find that the conditional effect of fluency on stock returns is positive and significant for stocks with a high sentiment beta, and not significant for stocks with a low sentiment beta. For stocks with a high sentiment beta, relative to stocks with a low sentiment beta, the effect of a one-point increment in the fluency score on excess stock returns increases by 0.13% with a one-standard-deviation increment in sentiment beta. For stocks with a plus one-standard-deviation sentiment beta, the effect of a one-point increment in the fluency score on excess stock returns is 0.04% (= −0.04%−0.02%+0.13%−0.03%). The results indicate that the positive interaction between sentiment beta and fluency from column (2) is entirely driven by stocks with a high sentiment beta. For stocks with a low sentiment beta, there is no conditional relation between fluency and returns.

Despite the large vector of controls, one potential concern is that the empirical model may not fully account for systematic risk. To address this issue, we augment the test equation with the factor loadings from the time-series analysis. Specifically, we estimate them individually for each stock in the sample using three-year rolling regressions as in Hong and Kacperczyk (2009). The results, reported in column (4), are robust to these additional controls. The inclusion of these known determinants of stock returns also leads to a substantial improvement of the goodness of fit, which increases from 9.8% to 50.7%.

Another potential concern with these results is that the return differential between fluent and nonfluent stocks might actually be driven by the overpricing of nonfluent stocks, rather than the underpricing of fluent stocks. We address this issue by breaking down the fluency index into separate fluency scores. Specifically, we introduce a set of dummy variables that take on value one for a given fluency score, and zero otherwise. Reassuringly, we find that the effect of fluency on stock returns from Table 4, column (4), is entirely confined to fluent stocks (scores two, three, and four), with the magnitude progressively increasing with the level of fluency, whereas the coefficient is close to zero for nonfluent stocks (scores zero and one).[17] The findings then lend support to the model prediction that unsophisticated investors undervalue fluent stocks.

Under the information hypothesis, the return differential between fluent and nonfluent stocks should increase with the quality difference among firms. To test this conjecture, we introduce a dummy variable that takes on value one for firms that have an above-median market capitalization on December 31 of the previous calendar year, and zero otherwise, and introduce interaction terms with the fluency and sentiment variables. The advantage of using a dummy is that the coefficient of categorical variables can be interpreted as abnormal returns in Fama–MacBeth regressions of this sort (Gompers et al., 2003; Mueller et al., 2017).[18]

The results are in Table 5. In column (1), we start the analysis again without conditioning on sentiment. We find that the coefficients of the standalone fluency variable and its interaction term with the size dummy are close to zero and not significant.

17 Due to the unusually large size of this table, we omit it for brevity. The results are available upon request.

18 This specification is then similar in spirit to the double-sorting on fluency and size from the portfolio analysis.

Table 5
Fama–MacBeth regressions of returns, fluency scores, and investor sentiment: Size breakdown.

Dep. Var.: Ri-Rf	(1) Full	(2) Full	(3) Full	(4) Excl. Micr.	(5) Excl. Penny
Fluency	−0.0003 (−0.72)	−0.0002 (−0.38)	−0.0003 (−0.95)	−0.0004 (−1.04)	−0.0003 (−0.96)
Large	−0.0009 (−0.88)	−0.0004 (−0.36)	−0.0024*** (−2.87)	−0.0027*** (−2.79)	−0.0030*** (−3.21)
Fluency × Large	0.0001 (0.24)	−0.0001 (−0.22)	0.0000 (0.12)	0.0001 (0.22)	0.0001 (0.23)
Fluency × Sent. Beta		0.0014*** (2.76)	0.0015*** (4.53)	0.0015*** (3.70)	0.0013*** (3.79)
Fluency × Sent. Beta × Large		−0.0013** (−2.14)	−0.0016*** (−4.44)	−0.0017*** (−3.83)	−0.0014*** (−3.58)
Sentiment Beta		−0.0039** (−2.59)	−0.0036*** (−2.67)	−0.0033** (−2.34)	−0.0034*** (−2.61)
Sentiment Beta × Large		0.0032** (2.41)	0.0042*** (4.34)	0.0042*** (3.80)	0.0038*** (4.15)
Book-to-Market (−1)	0.0015** (2.49)	0.0013** (2.19)	−0.0012** (−2.09)	−0.0011** (−2.10)	−0.0013** (−2.26)
Dividend Yield (−1)	0.0007 (1.49)	0.0007 (1.53)	−0.0004 (−1.24)	−0.0004 (−1.37)	−0.0002 (−0.68)
CumRet (−2,−3)	0.0012 (0.31)	0.0017 (0.46)	0.0032 (0.46)	0.0039 (0.56)	0.0027 (0.40)
CumRet (−4,−6)	0.0047 (1.31)	0.0067** (2.09)	0.0121** (2.44)	0.0121** (2.46)	0.0122** (2.43)
CumRet (−7,−12)	0.0104*** (3.63)	0.0114*** (4.39)	0.0198*** (6.67)	0.0193*** (6.48)	0.0190*** (6.55)
Price (−2)	−0.0017** (−2.25)	−0.0017** (−2.28)	0.0011*** (2.89)	0.0012*** (2.94)	0.0002 (0.50)
Volume (−2)	−0.0002 (−0.64)	−0.0003 (−0.93)	−0.0004* (−1.90)	−0.0004* (−1.89)	−0.0003 (−1.43)
Factor Loadings	N	N	Y	Y	Y
Observations	355,607	355,607	354,499	333,557	348,886
R-squared	0.0643	0.0965	0.5080	0.5107	0.5093

Fama–MacBeth regressions of U.S. monthly excess stock returns on company name fluency scores, investor sentiment beta, a dummy variable that takes on value one for firms whose market capitalization lies above the median on December 31 of the previous calendar year, interaction terms between fluency, sentiment beta, and the dummy, a set of firm characteristics from Edmans (2011), and factor loadings for the market, size, book-to-market, investment, and profitability factors from Fama and French (2015), the momentum factor from Carhart (1997), the liquidity factor from Pástor and Stambaugh (2003), the size, management, and performance factors from Stambaugh and Yuan (2017), and the book-to-market and profitability factors from Novy-Marx (2013), estimated over a rolling three-year window as in Hong and Kacperczyk (2009). The fluency scores are from Green and Jame (2013), measured on December 31 of each calendar year, and take on integer values between 0 (least fluent) and 4 (most fluent). We calculate sentiment betas as the winsorized (at the 1% tails) OLS coefficient of rolling five-year regressions of monthly excess stock returns on Baker and Wurgler's 2007 investor sentiment index, orthogonalized to business cycle indicators, controlling for excess returns on the stock market portfolio. To ease the interpretation of the estimates, all sentiment betas are standardized by subtracting their mean and dividing by their standard deviation. In column (4), we exclude microcaps, defined as the stocks that lie at the bottom 20% of the size distribution. In column (5), we exclude penny stocks, defined as stocks whose price is below $5 dollars. Stock data is from CRSP, accounting data is from Compustat, and the investor sentiment index is from Jeffrey Wurgler's website. The sample period is from January 1982 to December 2008, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

In column (2), however, we find that for stocks that exhibit a one-standard-deviation sentiment beta, the effect of a one-point increment in the fluency score on excess stock returns increases by 0.14% per month for firms with below-median market capitalization, whereas the effect is indistinguishable from zero for firms that lie above the median.

In column (3), we find similar results when augmenting the empirical model with the vector of factor loadings introduced above. Finally, we address the concern that our results might be driven by stocks of extremely small size. To this end, we exclude microcaps in column (4), defined as the stocks that lie at the bottom 20% of the size distribution (see, e.g., Green and Jame (2013)), and penny stocks in column (5), defined as stocks whose price is below $5 (see, e.g., Kumar (2009), Bhootra (2011)). Reassuringly, we find that the results are virtually unchanged. Overall, the findings mirror those from the time-series analysis, and lend support to the information hypothesis.

Table 6
Panel regressions of returns, fluency scores, and investor sentiment.

Dep. Var.: Ri-Rf	(1)	(2)	(3)	(4)	(5)
Fluency	0.0003	0.0004	−0.0001	−0.0004	−0.0005
	(0.46)	(0.60)	(−0.10)	(−0.34)	(−0.41)
Sentiment Beta		−0.0018**	−0.0016		−0.0060***
		(−2.20)	(−1.50)		(−3.30)
Fluency × Sent. Beta		0.0004	−0.0004		0.0015*
		(1.08)	(−0.74)		(1.95)
Fluency × Sent. Beta × High Beta			0.0013***		
			(3.56)		
High Beta			−0.0005		
			(−0.35)		
Fluency × High Beta			0.0001		
			(0.12)		
Large				−0.0106***	−0.0110***
				(−4.23)	(−4.42)
Fluency × Large				0.0010	0.0011
				(0.93)	(1.02)
Fluency × Sent. Beta × Large					−0.0016*
					(−1.86)
Sentiment Beta × Large					0.0056***
					(2.84)
Firm FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y
Observations	355,607	355,607	355,607	355,607	355,607
R-squared	0.0123	0.0124	0.0124	0.0101	0.0102

Panel regressions of U.S. monthly excess stock returns on company name fluency scores, investor sentiment beta, a dummy variable that takes on value one if the investor sentiment beta is positive, a dummy variable that takes on value one for firms whose market capitalization lies above the median on December 31 of the previous calendar year, interaction terms between fluency, sentiment beta, and the two dummy variables, and a set of controls from Edmans (2011). The fluency scores are from Green and Jame (2013), measured on December 31 of each calendar year, and take on integer values between 0 (least fluent) and 4 (most fluent). We calculate sentiment betas as the OLS coefficient of rolling five-year regressions of monthly excess stock returns on Baker and Wurgler's 2007 investor sentiment index, orthogonalized to business cycle indicators, controlling for excess returns on the stock market portfolio. We winsorize the 1% tails of the sentiment beta distribution. To ease the interpretation of the estimates, sentiment betas are standardized by subtracting their mean and dividing by their standard deviation. All specifications include firm and year fixed-effects. Standard errors are robust and clustered by firm. Stock data is from CRSP, accounting data is from Compustat, and the investor sentiment index is from Jeffrey Wurgler's website. The sample period is from January 1982 to December 2008, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

4.3. Fixed-effects regressions

Petersen (2009) shows that Fama–MacBeth regressions with highly persistent variables may generate biased standard errors. To address this concern, we repeat the analysis by estimating panel regressions with firm and year fixed-effects, and clustering standard errors by firm. This specification also addresses the concern that the between-firm estimates could be driven by time-invariant firm characteristics.

The results, reported in Table 6, are similar to those from the Fama–MacBeth regressions.[19] In column (1), we find that fluency as a standalone variable has no significant effect on stock returns. In column (2), we introduce the interaction terms with sentiment betas. The coefficient of the interaction term between fluency and sentiment beta is positive but outside the rejection region. In column (3), we show that this result varies dramatically across stocks with high and low sentiment betas. For stocks with a one-standard-deviation sentiment beta, the effect of a one-point increment in the fluency score on excess stock returns increases by 0.13%. The effect is not significant for stocks with a low sentiment beta.

In column (4), we find that the standalone fluency variable and its interaction term with the size dummy have no explanatory power over stock returns. In column (5), we condition the analysis on sentiment betas. We find that for stocks with a one-standard-deviation sentiment beta, the effect of a one-point increment in the fluency score on excess stock returns increases by 0.15% per month for firms with below-median market capitalization, whereas the effect is indistinguishable from zero for firms that lie above the median.[20]

[19] Similarly, Green and Jame (2013) find that the relation between fluency and firm valuations is robust to firm fixed-effects. Conversely, Karpoff and Rankine (1994) find that name changes are not associated with higher valuations, but they do not consider fluency in their analysis.

[20] We also find that the results are not specific to foreign sounding firms (defined as those with a low Englishness score), tech firms (defined as in Kile

Table 7
Fama–MacBeth and panel regressions of returns: Individual components fluency index.

	Fama-MacBeth			Fixed-Effects		
Dep. Var.: Ri-Rf	(1) Length	(2) Dictionary	(3) Englishness	(4) Length	(5) Dictionary	(6) Englishness
Fluency	−0.0009** (−1.98)	−0.0013** (−2.45)	−0.0006 (−0.88)	−0.0010 (−1.11)	−0.0006 (−0.34)	0.0008 (0.42)
Sentiment Beta	−0.0002 (−0.12)	−0.0009 (−0.88)	−0.0015 (−1.32)	−0.0010 (−1.19)	−0.0009* (−1.77)	−0.0016* (−1.68)
Fluency × Sent. Beta	−0.0010** (−2.08)	−0.0019** (−2.08)	−0.0005 (−0.54)	−0.0008 (−1.23)	−0.0028** (−2.12)	−0.0009 (−0.76)
Fluency × Sent. Beta × High Beta	0.0016** (2.23)	0.0058*** (5.46)	0.0033*** (3.63)	0.0017*** (2.64)	0.0055*** (3.24)	0.0034*** (3.50)
High Beta	−0.0011 (−1.23)	−0.0003 (−0.47)	−0.0001 (−0.06)	−0.0008 (−0.81)	−0.0004 (−0.63)	−0.0001 (−0.05)
Fluency × High Beta	0.0004 (0.54)	−0.0012 (−1.49)	−0.0007 (−0.63)	0.0004 (0.51)	0.0001 (0.08)	−0.0004 (−0.26)
Firm FE	N	N	N	Y	Y	Y
Year FE	N	N	N	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Observations	355,663	355,663	355,663	355,663	355,663	355,663
R-squared	0.0983	0.0975	0.0975	0.0372	0.0372	0.0373

Fama–MacBeth regressions (columns (1) to (3)) and panel regressions (column (4) to (6)) of U.S. monthly excess stock returns on company name individual fluency scores, investor sentiment beta, a dummy variable that takes on value one if the investor sentiment beta is positive, interaction terms between fluency, sentiment beta, and the dummy, and a set of firm characteristics from Edmans (2011). We identify fluency using the individual components of the company name fluency index from Green and Jame (2013), which includes a separate score for word length (columns (1) and (4)), dictionary (columns (2) and (5)), and Englishness (columns (3) and (6)). These scores are measured on December 31 of each calendar year. We calculate sentiment betas as the winsorized (at the 1% tails) OLS coefficient of rolling five-year regressions of monthly excess stock returns on Baker and Wurgler's 2007 investor sentiment index, orthogonalized to business cycle indicators, controlling for excess returns on the stock market portfolio. To ease the interpretation of the estimates, all sentiment betas are standardized by subtracting their mean and dividing by their standard deviation. The panel regressions include firm and year fixed-effects. Stock data is from CRSP, accounting data is from Compustat, and the investor sentiment index is from Jeffrey Wurgler's website. The sample period is from January 1982 to December 2008, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

4.4. Fluency index breakdown

In the last part of our analysis of returns, we look deeper into our results by performing a breakdown of the fluency index. We alternatively replace the overall index with each of its three individual components (length, dictionary, and Englishness), and separately re-estimate our equations of interest to assess the relative contribution of these fluency dimensions to our previous set of results.

The estimates are in Table 7. In columns (1) to (3), we re-estimate our baseline Fama–MacBeth regressions. We find again that the conditional positive effect of fluency on stock returns is confined to stocks with a high sentiment beta. For these stocks, relative to stocks with a low sentiment beta, the effect of a one-point increase in the length score on excess stock returns increases by 0.16% with a one-standard-deviation increment in sentiment beta. The overall effect of a one-point increment in the length score on the returns on stocks with a plus one-standard-deviation sentiment beta is 0.03% (= −0.09% − 0.10% + 0.16% + 0.04%). The effect is 0.14% for the dictionary score, and 0.15% for the Englishness score. The results are similar in columns (4) to (6), where we re-estimate our fixed-effects regressions.

Overall, then, all three components of the fluency index seem important in explaining the relation between fluency and stock returns.

and Phillips (2009), or firms with high idiosyncratic volatility (defined as in Ang et al. (2006, 2009)).

4.5. Valuations

Under the information hypothesis, the price difference between fluent and nonfluent stocks decreases with naive investor demand, and with the quality difference between fluent and nonfluent firms. Under the affect hypothesis, on the other hand, the price of fluent stocks is higher than that of nonfluent stocks. The price differential is proportional to the size of naive investor demand and the fluency bias. To test these conjectures, we estimate Fama–MacBeth valuation regressions from Hong and Kacperczyk (2009), and augment them with our variables of interest. We introduce sentiment betas in the analysis, defined as above, and then transformed into annual averages. We introduce interaction terms between fluency, sentiment betas, and size, and define company valuations as either Tobin's q or the market-to-book ratio.

Table 8 shows that the coefficient of fluency as a standalone variable is positive and highly significant, and the coefficient of the interaction term between fluency and size is negative and highly significant. The estimates imply that for a company at the 25th percentile of market capitalization, a one-point increase in fluency is associated with an increase in Tobin's q of 2.8%, whereas for a company that lies at the 75th percentile of market capitalization the effect drops to only 0.2%. The magnitudes are similar for the market-to-book ratio (4.2% and 0.1%, respectively). This empirical pattern supports the information story, as the effect of fluency on valuations is indeed weaker for stocks with above-median market capitalization.[21]

[21] Green and Jame (2013) also find that the price differential between fluent and nonfluent stocks is lower among large stocks. They show that a one unit

Table 8
Fluency, size, and valuations.

	Tobin's q		Market-to-Book	
	(1)	(2)	(3)	(4)
Fluency	0.0089***	0.0721***	0.0140***	0.1103***
	(3.85)	(7.38)	(5.21)	(10.49)
Fluency × Size		−0.0093***		−0.0145***
		(−6.39)		(−8.43)
Size	0.0367***	0.0576***	0.0744***	0.1071***
	(4.12)	(6.15)	(4.87)	(7.72)
ROE	0.4202***	0.4211***	0.9031***	0.9053***
	(4.08)	(4.09)	(5.61)	(5.62)
R&D/Sales	3.2931***	3.2660***	3.7963***	3.7525***
	(17.18)	(16.99)	(10.83)	(10.64)
R&D Missing	−0.0832	−0.2005***	−0.2868***	−0.4425***
	(−1.58)	(−4.17)	(−3.86)	(−5.21)
S&P 500	0.1012	0.1016	0.0987	0.0996
	(1.62)	(1.59)	(1.29)	(1.28)
ROE (+1)	0.1702**	0.1681**	0.3626***	0.3584***
	(2.18)	(2.17)	(3.05)	(3.04)
ROE (+2)	0.0738	0.0741	0.1100	0.1103
	(0.88)	(0.89)	(0.86)	(0.86)
ROE (+3)	−0.0806**	−0.0821**	−0.0663	−0.0681
	(−2.08)	(−2.12)	(−1.15)	(−1.18)
Fama-MacBeth	Y	Y	Y	Y
Observations	10,465	10,465	10,465	10,465
R-squared	0.2575	0.2599	0.4246	0.4272

Fama–MacBeth regressions of the natural logarithm of Tobin's q (columns 1 and 2), defined as the ratio between enterprise value (debt plus market value of equity) and book value (debt plus book value of equity), or the market-to-book ratio (columns 3 and 4), defined as market capitalization divided by the book value of equity, on the company name fluency score, company size, defined as the natural logarithm of market capitalization, an interaction term between fluency and size, and a vector of controls including the fraction of research and development expenditures to firm sales, a dummy variable that takes on value one if the company's R&D expenditure is missing in a given year, a dummy variable that takes on value one if the company is part of the S&P 500 index in a given year, and return on equity, calculated over the subsequent three years, and defined as the ratio of earnings over the book value of equity, where earnings are calculated as income before extraordinary items, plus deferred taxes, plus investment tax credit. The fluency scores are from Green and Jame (2013), measured on December 31 of the previous calendar year, and take on integer values between 0 (least fluent) and 4 (most fluent). Observations are annual. We winsorize the 1% tails of the distribution of Tobin's q, market-to-book, and size. Stock data is from CRSP, and accounting data is from Compustat. The sample period is from 1981 to 2007, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table 9 shows that the coefficient of the interaction term between fluency and sentiment is negative and highly significant in all specifications. To get a sense of the magnitude, consider a firm of median size (i.e., with a market cap of 435 USD million). If the stock's sentiment beta is zero, a one-point increase in the fluency index is associated with an increase in Tobin's q of 2.0%. If the stock exhibits a one-standard-deviation sentiment beta, the effect drops to 1.6%. For the market-to-book ratio, the magnitudes are 2.4% and 2.1%, respectively. We obtain similar results when excluding microcaps and penny stocks from the sample.

increase in fluency is associated with a $1.5 million increase in market equity for microcap stocks, and $10 million for non-microcap stocks (see Green and Jame (2013, page 824, Section 6.2). In this paper, we complement their subsample breakdown with a parametric analysis.

The estimates are in line with the idea that a large participation of naive investors brings the valuations of fluent and nonfluent stocks closer to each other, as naive investors fail to recognize the information content of fluency. The negative coefficient for the interaction term between sentiment and fluency is inconsistent with the affect hypothesis. The empirical evidence on valuations then lends strong support to the information hypothesis, much like the analysis of returns.

4.6. Additional tests

In his 2017 AFA Presidential Address, Campbell R. Harvey tells a valuable cautionary tale on how research in asset pricing can produce spurious results in the absence of a clear economic mechanism. To address this concern, we test the two underpinnings of the information story.

First, we analyze whether fluency is indeed informative, i.e., whether it identifies superior firms in terms of their future operating performance. This is a crucial test, as in the affect story there is no correlation between fluency and the quality of the firm. To this end, we analyze the relation between return on assets, defined as either EBITDA or EBIT divided by the book value of assets, and beginning-of-year company name fluency. We estimate Fama–MacBeth regressions using a set of firm-level controls from Hong and Kacperczyk (2009), and also introduce an interaction term between fluency and size.[22]

The results are in Table 10, columns (1) and (2). We find that the coefficient of the fluency index is associated with an increase in return on assets, and the effect again decreases with firm size. In either specification, the effect of a one-point increase in fluency on ROA is 6% higher for a company at the 25th percentile of market capitalization if compared with a company at the 75th percentile of market capitalization. The magnitude of the effect is similar when we measure ROA over the subsequent three years, in columns (3) and (4), which allays the concern that accounting data is released with some delay.

Second, we test the assumption that the mispricing of fluency comes from expectation errors. Following Engelberg et al. (2018), we conjecture that this bias, if present among naive investors, should also characterize the forecasts of less sophisticated analysts. The mechanism we hypothesize is as follows. If some analysts neglect the information about good fundamentals captured by fluency, their expectations should be systematically too pessimistic for fluent firms. When aggregating forecasts across all analysts, then, fluent firms should produce positive earnings surprises. To test for this, we estimate the fixed-effects earnings surprise regressions from Mueller et al. (2017). Again, we augment the test equation with an interaction term between fluency and size.

The results are in Table 11, columns (1) and (2). We find a positive association between fluency and earnings surprises, and again the effect decreases with size. For a company at the 25th percentile of market capitalization, the effect of a one-point increase in fluency on earnings surprises is higher than for a company at the 75th percentile of market capitalization by 1.1% of the beginning-of-period stock price. The magnitude is similar when re-estimating the regressions using the Fama–MacBeth procedure (0.9%). The estimates rise to 1.4% and 1.3%, respectively, when we measure earnings surprises over the subsequent three quarters, in columns (3) and (4). The empirical evidence lends support to our conjecture that the mispricing of fluency comes from expectation errors.

[22] Here we leave out sentiment betas, because they have no impact on company performance.

Table 9
Fluency, sentiment, size, and valuations.

	Tobin's q			Market-to-Book		
	(1) Full	(2) Excl. Micr.	(3) Excl. Penny	(4) Full	(5) Excl. Micr.	(6) Excl. Penny
Fluency	0.0750*** (5.77)	0.0929*** (4.91)	0.0738*** (5.21)	0.1125*** (9.87)	0.1674*** (6.91)	0.1111*** (7.66)
Fluency × Size	−0.0091*** (−5.20)	−0.0114*** (−4.77)	−0.0089*** (−4.82)	−0.0146*** (−9.81)	−0.0219*** (−7.01)	−0.0144*** (−8.06)
Size	0.0613*** (5.39)	0.0562*** (4.66)	0.0607*** (5.39)	0.1104*** (6.83)	0.1116*** (7.16)	0.1101*** (7.03)
Sentiment Beta	0.1294** (2.65)	0.2345*** (2.97)	0.1281** (2.37)	0.1349** (2.28)	0.1877* (1.99)	0.1307* (2.00)
Fluency × Sent. Beta	−0.0642*** (−4.40)	−0.1077*** (−4.01)	−0.0667*** (−3.87)	−0.0619*** (−3.62)	−0.0782** (−2.31)	−0.0641*** (−2.99)
Fluency × Sent. Beta × Size	0.0099*** (5.11)	0.0152*** (4.14)	0.0101*** (4.63)	0.0097*** (3.73)	0.0109** (2.18)	0.0097*** (3.37)
Size × Sentiment Beta	−0.0126** (−2.16)	−0.0264** (−2.38)	−0.0121* (−1.91)	−0.0136* (−2.02)	−0.0199 (−1.44)	−0.0126* (−1.72)
ROE	0.4131*** (4.00)	0.3920*** (3.95)	0.4144*** (4.00)	0.8989*** (5.57)	0.8738*** (5.62)	0.9015*** (5.58)
R&D/Sales	3.0893*** (15.46)	3.0837*** (16.62)	3.1025*** (15.75)	3.5838*** (10.08)	3.5253*** (10.17)	3.5893*** (10.12)
R&D Missing	−0.2214*** (−3.68)	−0.1744*** (−3.61)	−0.2223*** (−4.08)	−0.5061*** (−5.93)	−0.4706*** (−4.58)	−0.3791*** (−5.28)
S&P 500	0.0965 (1.49)	0.0940 (1.44)	0.0957 (1.48)	0.0900 (1.16)	0.0879 (1.12)	0.0891 (1.15)
ROE (+1)	0.1774** (2.25)	0.1973** (2.54)	0.1737** (2.19)	0.3622*** (3.08)	0.3917*** (3.26)	0.3598*** (3.02)
ROE (+2)	0.0777 (0.89)	0.0590 (0.70)	0.0766 (0.86)	0.1171 (0.86)	0.0957 (0.71)	0.1136 (0.82)
ROE (+3)	−0.0799* (−2.05)	−0.0683* (−1.85)	−0.0859** (−2.27)	−0.0688 (−1.18)	−0.0496 (−0.89)	−0.0772 (−1.38)
Fama-MacBeth	Y	Y	Y	Y	Y	Y
Observations	10,465	9,233	10,325	10,465	9,233	10,325
R-squared	0.2817	0.2626	0.2802	0.4412	0.4201	0.4385

Fama–MacBeth regressions of the natural logarithm of Tobin's q (columns 1 and 2), defined as the ratio between enterprise value (debt plus market value of equity) and book value (debt plus book value of equity), or the market-to-book ratio (columns 3 and 4), defined as market capitalization divided by the book value of equity, on the company name fluency score, investor sentiment beta, company size, defined as the natural logarithm of market capitalization, interaction terms between fluency, sentiment beta, and size, and a vector of controls including the fraction of research and development expenditures to firm sales, a dummy variable that takes on value one if the company's R&D expenditure is missing in a given year, a dummy variable that takes on value one if the company is part of the S&P 500 index in a given year, and return on equity, calculated over the subsequent three years, and defined as the ratio of earnings over the book value of equity, where earnings are calculated as income before extraordinary items, plus deferred taxes, plus investment tax credit. The fluency scores are from Green and Jame (2013), measured on December 31 of the previous calendar year, and take on integer values between 0 (least fluent) and 4 (most fluent). Observations are annual. We winsorize the 1% tails of the distribution of Tobin's q, market-to-book, and size. Sentiment betas are calculated as the OLS coefficient of rolling five-year regressions of monthly excess stock returns on Baker and Wurgler's 2007 investor sentiment index, orthogonalized to business cycle indicators, controlling for excess returns on the stock market portfolio, and then transformed into annual averages. To ease the interpretation of the estimates, all sentiment betas are standardized by subtracting their mean and dividing by their standard deviation. In columns (2) and (5), we exclude microcaps, defined as the stocks that lie at the bottom 20% of the size distribution. In column (3) and (6), we exclude penny stocks, defined as stocks whose price is below $5 dollars. Stock data is from CRSP, accounting data is from Compustat, and the investor sentiment index is from Jeffrey Wurgler's website. The sample period is from 1981 to 2007, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

In unreported tests, we also consider two alternative specifications of the fluency index. First, we create a restricted version of the index as in Green and Jame (2013), grouping together the least fluent scores (0 and 1) and the most fluent scores (3 and 4), respectively. The restricted fluency index then takes on integer values between one and three, and addresses the concern that the extreme fluency categories (zero and four) exhibit a relatively small number of firm-year observations. Second, we construct a dummy variable that takes on value one if the fluency index takes on the most fluent scores (3 and 4), and zero otherwise. In the spirit of Mueller et al. (2017) we exclude company names with a middle score (2) when using the fluency dummy, so that we can effectively measure the difference in returns between the top two and the bottom two fluency categories. We find that our entire set of empirical results is robust to these alternative specifications.

4.7. Extended sample

All previous analyses consider the same fluency scores and sample period as in Green and Jame (2013). The advantage of

Table 10
Fluency and operating performance.

	ROA (+1)		ROA (+2,+4)	
	(1) EBITDA	(2) EBIT	(3) EBITDA	(4) EBIT
Fluency	0.2147*** (6.24)	0.1954*** (6.68)	0.2173*** (6.39)	0.2044*** (6.72)
Fluency × Size	−0.0206*** (−5.51)	−0.0203*** (−6.56)	−0.0222*** (−6.00)	−0.0223*** (−6.61)
Size	0.0463*** (3.58)	0.0328** (2.72)	0.0527*** (3.89)	0.0386*** (2.90)
Book-to-Market	−0.5460*** (−32.84)	−0.6525*** (−43.82)	−0.5002*** (−26.27)	−0.5740*** (−34.33)
Price Inverse	1.0673*** (7.61)	0.7373*** (5.52)	2.0578*** (8.58)	1.9917*** (9.42)
St. Dev. Returns	2.0368*** (3.88)	0.6424 (1.14)	2.4582*** (4.08)	1.5683** (2.59)
Market Beta	0.0470* (1.79)	0.0514** (2.21)	0.0149 (0.55)	0.0247 (0.95)
Stock Returns	0.0846 (0.16)	1.9103*** (3.90)	−1.2129* (−1.90)	−0.3058 (−0.56)
Nasdaq	−0.3430*** (−7.95)	−0.2558*** (−6.73)	−0.3572*** (−7.43)	−0.2936*** (−6.57)
S&P 500	−0.0437 (−0.69)	−0.0118 (−0.21)	−0.0386 (−0.58)	−0.0197 (−0.36)
Fama-MacBeth	Y	Y	Y	Y
Observations	36,115	36,115	24,744	24,744
R-squared	0.3367	0.3428	0.3224	0.3423

Fama–MacBeth regressions of return on assets (ROA), defined as EBITDA (columns 1 and 3) or EBIT (columns 2 and 4) divided by the book value of assets, on company name fluency, company size, defined as the natural logarithm of market capitalization, an interaction term between fluency and size, the natural logarithm of the ratio between market capitalization and book value, the inverse of the share price at the end of the year, the standard deviation of the residuals of the company's daily stock returns from Carhart's (1997) four-factor model over a given year, the company's market beta, the average monthly return during a given year, a dummy variable that takes on value one if the stock is listed on Nasdaq and zero otherwise, and a dummy variable that takes on value one if the stock is part of the S&P 500 index and zero otherwise. The fluency scores are from Green and Jame (2013), measured on December 31 of the previous fiscal year, and take on integer values between 0 (least fluent) and 4 (most fluent). We winsorize the 1% tails of the distribution of ROA and size. In columns (1) and (2), ROA is measured over the next year. In columns (3) and (4), we consider the average ROA over the subsequent three years, i.e., two to four years ahead. Accounting data is from Compustat, and the investor sentiment index is from Jeffrey Wurgler's website. The sample period is from 1981 to 2007, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

this approach is that it makes our results directly comparable to theirs. However, it is interesting to analyze whether our findings also hold in more recent times. To address this point, we construct an updated version of the fluency index that extends until 2021, and repeat our main analyses of returns in the extended sample.

To obtain our updated fluency measures, we follow the same procedure as in Green and Jame (2013).[23] We use the most recent vintages of the Microsoft Office 365 spell-check (2023) and the top 60,000 lemmas from the Corpus of Contemporary American English (2021) to construct the dictionary and Englishness measures. For the length measure, we use their exact same algorithm.

The results are in Table 12. In Panel A, we estimate Fama–MacBeth regressions. Column (1) considers the overall fluency index. As in the original sample from Green and Jame (2013), we find the main prediction of the model confirmed: the coefficient of the triple interaction term between fluency, sentiment beta, and the positive-beta dummy is positive and significant, and equal to 0.17%. This effect size is comparable to its counterpart from Table 4 (0.13%). The overall effect of fluency, however, decreases substantially. Following a one-point increase in the fluency index in the extended sample, the returns on stocks with a plus one-standard-deviation sentiment beta now decrease by 0.05% (= $-0.06\% - 0.19\% + 0.17\% + 0.03\%$).

To shed further light on this finding, we decompose the fluency index into its three individual components as in Table 7. The results are in Table 12, columns (2) to (4). The coefficient of the triple interaction term between fluency, sentiment beta, and the positive-beta dummy is again positive and significant for all three specifications. However, the overall effect of fluency varies substantially across individual measures. For length fluency, a one-point increment leads to an increase in returns on stocks with a plus one-standard-deviation sentiment beta of 0.03% (= $-0.07\%-0.18\%+0.26\%+0.02\%$), which mirrors the point estimate from the original sample. For dictionary and Englishness, on the other hand, the total effect becomes negative and equal to −0.11% and −0.12%, respectively.

In Panel B, we repeat the analysis by estimating fixed-effects regressions. We find similar results and an even more clear-cut empirical pattern. The coefficient of the triple interaction term of interest is significant for length fluency only (0.15%), implying that a one-point increase in length fluency is associated with an increase in returns on stocks with a plus one-standard-deviation sentiment beta of 0.03% (= $-0.06\%-0.10\%+0.15\%+0.04\%$). This estimate is identical to its counterpart from the Fama–MacBeth regressions from Panel A. Conversely, all the coefficients of the fluency variables lie outside the rejection region when using the dictionary and Englishness scores.

These results for the extended sample period differ from the fluency breakdown performed in Table 7. In the original sample from Green and Jame (2013), all three fluency measures have a positive and significant effect on the returns on stocks with a high sentiment beta. In the extended sample, only length fluency retains its explanatory power. The effects for the dictionary and Englishness measures become weaker.

This sharp contrast might reflect two drawbacks of the dictionary and Englishness scores in the extended sample. First, they only represent the most recent state of the English language. Their scores respectively depend on the particular vintage of Microsoft spell-check and the Corpus under consideration. Since the analysis considers the latest vintages available, these measures include some potential measurement error due to the application of recent language rules to a sample period spanning four decades over which language has likely evolved.

This drawback is further compounded by a second kind of measurement error. Neither the dictionary nor the Englishness measure include data from social media among their sources, which might be an important omission because social media constitutes one of the main drivers of modern language change.[24] Therefore, there are companies whose names receive low dictionary and/or Englishness scores whereas such names would generally be considered fluent nowadays (e.g., "Xtra" as a phonetic

[23] The procedure is described in greater detail in the working paper version of their article, available on SSRN.

[24] See, e.g., "Research shows Twitter is driving English language evolution", Brandwatch, May 29, 2013; or "ICYMI, English language is changing faster than ever", the Guardian, May 1, 2015.

Table 11
Fluency and earnings forecasts.

	Earnings Surprise (+1) / Stock Price		Earnings Surprise (+2,+4) / Stock Price	
	(1)	(2)	(3)	(4)
Fluency	0.0479***	0.0348***	0.0628***	0.0576***
	(6.57)	(5.42)	(7.10)	(9.53)
Fluency × Size	−0.0039***	−0.0031***	−0.0049***	−0.0046***
	(−7.17)	(−6.42)	(−7.38)	(−10.38)
Size	0.0116***	0.0094***	0.0148***	0.0136***
	(5.76)	(5.84)	(6.58)	(9.58)
Book-to-Market	0.0059***	0.0059***	0.0048**	0.0049***
	(3.09)	(3.94)	(2.34)	(3.29)
Year FE	Y	N	Y	N
Quarter FE	Y	N	Y	N
Fama-MacBeth	N	Y	N	Y
Observations	2,409	2,409	2,055	2,055
R-squared	0.1127	0.2409	0.1917	0.3040

Panel regressions of quarterly analysts' forecast error (earnings surprises), defined as the firm's actual earnings per share at the end of the quarter minus the I/B/E/S consensus forecast of earnings per share, scaled down by the company's stock price at the beginning of the quarter, on company name fluency, company size, defined as the natural logarithm of market capitalization, an interaction term between fluency and size, and the natural logarithm of the ratio between market capitalization and book value. The fluency scores are from Green and Jame (2013), measured on December 31 of the previous fiscal year, and take on integer values between 0 (least fluent) and 4 (most fluent). In columns (3) and (4), we consider the average earnings surprise over the subsequent three quarters, i.e., two to four quarters ahead. In columns (1) and (3), the regressions include quarter and year fixed-effects, with robust standard errors clustered at both the firm and year level. In columns (2) and (4), the regressions are estimated through the Fama–MacBeth procedure. Stock data is from CRSP, while analyst forecast data is from I/B/E/S. The sample period is from the first quarter of 1992 to the fourth quarter of 2008, and t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

spelling of the word "extra"; the latter passes the spell-check, but the former does not).[25]

In unreported analyses, we also perform two groups of additional tests. First, we test our conjecture that the results from the returns regressions should be stronger for companies of small size. However, the coefficients of interest are outside of the rejection region. Second, we repeat our valuation tests in the extended sample. While we find again that fluent firms exhibit significantly higher valuations, the coefficients of the interaction terms with firm size are outside of the rejection region. The main hurdle of these analyses is that the correlation between company name fluency and firm size jumps up in recent times, with an exceptionally strong increase in market capitalization for highly-fluent companies.[26] As a result, the analysis of size as a moderating variable likely introduces collinearity issues.

5. Conclusion

Research shows that stocks with fluent names exhibit higher valuations. In this paper, we tease out the two competing explanations for this result. In the affect hypothesis, naive investors erroneously believe that fluency identifies better stocks. Their overbidding for fluent stocks generates overpricing, and subsequent lower returns. Under the information hypothesis, fluency actually correlates with the future profitability of the firm. If naive investors do not recognize the information embedded in fluency, fluent stocks become underpriced and yield higher returns.

While our results cannot definitively rule out the affect hypothesis, they suggest that the information hypothesis dominates. Fluent companies yield higher abnormal returns relative to non-fluent companies, exhibit superior future operating performance, and surprise analysts with positive unexpected earnings. Consistent with our theoretical model, these effects are concentrated among firms with low market capitalization and high sensitivity to investor sentiment. The identification of a specific mechanism through which fluency affects stock prices also addresses the concern that the results may be driven by spurious correlations. Altogether, the findings shed new light on the economic meaning of fluency.

Previous research shows that investors can misplace attention (see, e.g., Hirshleifer and Teoh (2003)). Specifically, they can either pay attention to something that should be ignored, or fail to attend to something that should be taken into account. In the context of the analysis of company names, Cooper et al. (2001) find evidence for mistakes of the first type. They show that unsophisticated investors react to irrelevant features of company names, such as a mere association with the Internet, leading to a large value increase for the firm. In this paper, we find that evidence for the second type of mistake. We show that unsophisticated investors ignore relevant information contained in the fluency of company names.

The present work also opens at least three potential avenues for future research. First, the exact mechanism behind the relation between fluency and firm quality has remained largely unresolved. One possibility is that a fluent name conveys a company's inner identity or culture to stakeholders, which in turn increases its human intellectual capital. Another possibility is that consumers have an affect-driven preference for products of fluently-named companies, boosting sales and operating performance. Fluent names may increase sales and margins also through improved brand recognition and media coverage. These effects might be strong especially for companies whose name coincides with a successful brand.

25 Attesting to this concern, we find that the original dictionary and Englishness measures from Green and Jame (2013) exhibit correlation coefficients of only 0.3 with the extended-sample counterparts, i.e., those obtained using the most recent data sources.

26 In simple pooled OLS regressions of market capitalization on the fluency index, the coefficient of interest is near-zero in magnitude and significance in the original sample from Green and Jame (2013) whereas it becomes positive, large, and highly significant in the more recent sample. In additional analyses, we find that this pattern seems mostly driven by the most fluent company names (score 4).

Table 12
Fluency and returns: Extended sample period.

Panel A. Fama-MacBeth

Dep. Var.: Ri-Rf	(1) Index	(2) Length	(3) Dictionary	(4) Englishness
Fluency	−0.0006**	−0.0007**	−0.0011***	−0.0013**
	(−2.58)	(−2.18)	(−2.60)	(−2.55)
Sentiment Beta	0.0068	0.0052	0.0055	0.0057
	(1.59)	(1.24)	(1.32)	(1.32)
Fluency × Sent. Beta	−0.0019***	−0.0018**	−0.0050**	−0.0036**
	(−2.92)	(−2.09)	(−2.52)	(−2.04)
Fluency × Sent. Beta × High Beta	0.0017**	0.0026**	0.0045**	0.0040**
	(2.46)	(2.45)	(2.08)	(2.11)
High Beta	−0.0003	0.0000	0.0000	−0.0002
	(−0.22)	(0.04)	(0.04)	(−0.19)
Fluency × High Beta	0.0003	0.0002	0.0005	0.0007
	(0.89)	(0.44)	(0.89)	(1.00)
Controls	Y	Y	Y	Y
Observations	670,359	670,359	670,359	670,359
R-squared	0.3506	0.3504	0.3502	0.3506

Panel B. Fixed-Effects

Dep. Var.: Ri-Rf	(1) Index	(2) Length	(3) Dictionary	(4) Englishness
Fluency	−0.0013	−0.0006	0.0014	−0.0098*
	(−0.59)	(−0.18)	(0.32)	(−1.66)
Sentiment Beta	0.0020***	0.0023***	0.0019***	0.0017**
	(2.67)	(4.50)	(5.10)	(2.35)
Fluency × Sent. Beta	−0.0003	−0.0010*	−0.0002	0.0001
	(−0.79)	(−1.93)	(−0.21)	(0.09)
Fluency × Sent. Beta × High Beta	0.0007**	0.0015***	0.0015	0.0008
	(2.18)	(2.67)	(1.26)	(1.04)
High Beta	−0.0011	−0.0008	−0.0003	−0.0010
	(−1.16)	(−1.30)	(−0.58)	(−1.13)
Fluency × High Beta	0.0003	0.0004	−0.0004	0.0007
	(0.81)	(0.89)	(−0.52)	(0.76)
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Observations	670,359	670,359	670,359	670,359
R-squared	0.0352	0.0352	0.0351	0.0351

Fama–MacBeth regressions (Panel A) and panel regressions (Panel B) of U.S. monthly excess stock returns on company name fluency scores, investor sentiment beta, a dummy variable that takes on value one if the investor sentiment beta is positive, interaction terms between fluency, sentiment beta, and the dummy, and a set of firm characteristics from Edmans (2011). The fluency scores are obtained following the procedure described in detail in the working paper version of Green and Jame (2013), and measured on December 31 of each calendar year. The length measure is built in the same way as the original one. The dictionary and Englishness measures are respectively constructed using the most recent vintages of the Microsoft Office 365 spell-check (2023) and the top 60,000 lemmas from the Corpus of Contemporary American English (2021). We include the full fluency index in column (1), and its breakdown in its three individual components (length, dictionary, Englishness) in columns (2) to (4). We calculate sentiment betas as the winsorized (at the 1% tails) OLS coefficient of rolling five-year regressions of monthly excess stock returns on Baker and Wurgler's 2007 investor sentiment index, orthogonalized to business cycle indicators, controlling for excess returns on the stock market portfolio. To ease the interpretation of the estimates, all sentiment betas are standardized by subtracting their mean and dividing by their standard deviation. The panel regressions include firm and year fixed-effects. Stock data is from CRSP, accounting data is from Compustat, and the investor sentiment index is from Jeffrey Wurgler's website. The sample period is from January 1982 to December 2021, and Newey–West t-statistics are reported in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Second, the analysis of the informational content of fluency can be extended to alternative asset classes, such as corporate bonds or equity derivatives. These markets may include a different proportion of sophisticated investors, which would affect the pricing of company name fluency.

Third, our study can be replicated for different countries. While it might be challenging to identify counterparts to the Englishness and spell-check measures for foreign languages, our analysis supports the view that name length, despite its coarse nature, might provide a relatively simple way to assess company name fluency in other countries (see, e.g., Jin et al. (2021)). In particular, it would be interesting to relate the mispricing of fluency to the country's degree of market efficiency.

CRediT authorship contribution statement

Maurizio Montone: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Validation. **Martijn J. van den Assem:** Conceptualization, Investigation, Methodology, Project administration, Writing – review & editing, Supervision. **Remco C.J. Zwinkels:** Conceptualization, Investigation, Methodology, Project administration, Writing – review & editing, Supervision.

Appendix

First-order condition

The first-order condition for security i is:

$$E_1^{\phi}\left(\sum_{i=1}^{n}(S_{2i}-S_{1i})\right)-\frac{\gamma}{2}\left(2x_i^{\phi}var_1^{\phi}(S_{2i})+\sum_{i\neq j}2x_j^{\phi}cov_1^{\phi}(S_{2i},S_{2j})\right)=0. \tag{A.1}$$

Solving out for the optimal investment:

$$x_i^{\phi}=\frac{E_1^{\phi}(S_{2i})-S_{1i}-z_{1ij}^{\phi}}{\gamma var_1^{\phi}(S_{2i})}, \tag{A.2}$$

where

$$z_{1ij}^{\phi}\equiv\gamma\sum_{i\neq j}x_j^{\phi}cov_1^{\phi}(S_{2i},S_{2j}). \tag{A.3}$$

Market clearing

For security i, the market-clearing condition is:

$$f\frac{E_1^N(S_{2i})-S_{1i}-z_{1ij}^N}{\gamma var_1^N(S_{2i})}+(1-f)\frac{E_1^A(S_{2i})-S_{1i}-z_{1ij}^A}{\gamma var_1^A(S_{2i})}=x_{0i}. \tag{A.4}$$

Solving out for the equilibrium price:

$$S_{1i}=\kappa_iE_1^N(S_{2i})+(1-\kappa_i)E_1^A(S_{2i})-\frac{\alpha_i^Nz_{1ij}^N+\alpha_i^Az_{1ij}^A}{\alpha_i^N+\alpha_i^A}-\frac{\gamma x_{0i}}{\alpha_i^N+\alpha_i^A}, \tag{A.5}$$

where:

$$\alpha_i^N\equiv\frac{f}{var_1^N(S_{2i})}, \tag{A.6}$$

$$\alpha_i^A\equiv\frac{1-f}{var_1^A(S_{2i})}, \tag{A.7}$$

$$\kappa_i\equiv\frac{\alpha_i^N}{\alpha_i^N+\alpha_i^A}. \tag{A.8}$$

For simplicity and without loss of generality, we assume that investors only disagree on the first moments of the return distribution.[27] We set:

$$cov_1^A(S_{2i},S_{2j})=cov_1^N(S_{2i},S_{2j})=cov_1(S_{2i},S_{2j}), \tag{A.9}$$

for all i, j. This immediately implies $\kappa_i = f$ for all i in the equilibrium price, and the last two addends can be rearranged as:

$$\frac{\alpha_i^Nz_{1ij}^N+\alpha_i^Az_{1ij}^A}{\alpha_i^N+\alpha_i^A}=\gamma\sum_{i\neq j}cov_1(S_{2i},S_{2j})(fx_j^N+(1-f)x_j^A), \tag{A.10}$$

$$\frac{\gamma x_{0i}}{\alpha_i^N+\alpha_i^A}=\gamma var_1(S_{2i})x_{0i}. \tag{A.11}$$

Also, note that the market-clearing condition implies:

$$fx_j^N+(1-f)x_j^A=x_{0j} \tag{A.12}$$

for all j. Then the equilibrium price can be expressed in closed form as a function of expectations and initial endowments:

$$S_{1i}=fE_1^N(S_{2i})+(1-f)E_1^A(S_{2i})-\gamma\sum_{i\neq j}cov_1(S_{2i},S_{2j})x_{0j}-\gamma var_1(S_{2i})x_{0i}. \tag{A.13}$$

Setting all securities in zero net supply, as in Hirshleifer and Teoh's (2003) original setting, yields the equilibrium price from Eq. (2), which generalizes Hirshleifer and Teoh's (2003) pricing equation to an economy with n risky securities.

Proof of Propositions 1 and 2

Under the affect hypothesis, the market clearing conditions for stocks with high and low fluency yield the following equilibrium prices:

$$S_1^H=1+f(b_H-1), \tag{A.14}$$

$$S_1^L=1-f(1-b_L), \tag{A.15}$$

which yields Proposition 1. Next, we define expected returns as the difference between the expected final cash flow and the stock price at time 1 (see, e.g., Chen, Hong, and Stein, 2002). In equilibrium, fluent stocks earn negative abnormal returns while nonfluent stocks earn positive abnormal returns:

$$E(\tilde{r}_2^H)=-f(b_H-1), \tag{A.16}$$

$$E(\tilde{r}_2^L)=f(1-b_L), \tag{A.17}$$

which implies Proposition 2.

Proof of Propositions 3 and 4

The information hypothesis yields the following equilibrium prices:

$$S_1^H=\lambda(f+(1-f)\mu), \tag{A.18}$$

$$S_1^L=\lambda. \tag{A.19}$$

27 See, e.g., Hong and Sraer (2016) for an excellent discussion on this point.

Abnormal returns are respectively:

$$E(\tilde{r}_2^H) = \lambda f(\mu - 1), \quad (A.20)$$

$$E(\tilde{r}_2^L) = 0. \quad (A.21)$$

Taking the price and return differentials yields Propositions 3 and 4.

References

Aaker, David A., 1996. Resisting temptations to change a brand position/execution: The power of consistency over time. J. Brand Manag. 3, 251–258.
Alter, Adam L., Oppenheimer, Daniel M., 2006. Predicting short-term stock fluctuations by using processing fluency. Proc. Natl. Acad. Sci. 103, 9369–9372.
Alter, Adam L., Oppenheimer, Daniel M., 2008. Easy on the mind, easy on the wallet: The roles of familiarity and processing fluency in valuation judgements. Psychonomic Bull. Rev. 15, 985–990.
Alter, Adam L., Oppenheimer, Daniel M., 2009. Uniting the tribes of fluency to form a metacognitive nation. Pers. Soc. Psychol. Rev. 13, 219–235.
Anderson, Alyssa G., Larkin, Yelena, 2019. Does non-informative text affect investor behavior? Financ. Manag. 48, 257–289.
Ang, Andrew, Hodrick, Robert J., Xing, Yuhang, Zhang, Xiaoyan, 2006. The cross-section of volatility and expected returns. J. Finance 61, 259–299.
Ang, Andrew, Hodrick, Robert J., Xing, Yuang, Zhang, Xiaoyan, 2009. High idiosyncratic volatility and low returns: International and further U.S. evidence. J. Financ. Econ. 91, 1–23.
Baker, Malcolm, Wurgler, Jeffrey, 2006. Investor sentiment and the cross-section of stock returns. J. Finance 61, 1645–1680.
Baker, Malcolm, Wurgler, Jeffrey, 2007. Investor sentiment in the stock market. J. Econ. Perspect. 21, 129–157.
Baker, Malcolm, Wurgler, Jeffrey, Yuan, Yu, 2012. Global, local, and contagious investor sentiment. J. Financ. Econ. 104, 272–287.
Barber, Brad, Odean, Terrance, 2002. Online investors: Do the slow die first? Rev. Financ. Stud. 15, 455–488.
Bart, Christopher K., 2001. Measuring the mission effect in human intellectual capital. J. Intellect. Cap. 2, 320–330.
Bhootra, Ajay, 2011. Are momentum profits driven by the cross-sectional dispersion in expected stock returns? J. Financ. Mark. 14, 494–513.
Bontis, Nick, 1999. Managing organizational knowledge by diagnosing intellectual capital: Framing and advancing the state of the field. Int. J. Technol. Manage. 18, 433–462.
Brennan, Michael, Chordia, Tarun, Subrahmanyam, Avanidhar, 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. J. Financ. Econ. 49, 345–373.
Carhart, Mark M., 1997. On persistence of mutual fund performance. J. Finance 52, 57–82.
Chan, Louis K.C., Lakonishok, Josef, Sougiannis, Theodore, 2001. The stock market valuation of research and development expenditures. J. Finance 56, 2431–2456.
Chan, Chien-Sheng Richard, Park, Haemin Dennis, Patel, Pankaj, 2018. The effect of company name fluency on venture investment decisions and IPO underpricing. Venture Capital 20, 1–26.
Chen, Joseph, Hong, Harrison, Stein, Jeremy C., 2002. Breadth of ownership and stock returns. J. Financ. Econ. 66, 171–205.
Cooper, Michael J., Dimitrov, Orlin, Rau, Raghavendra P., 2001. A rose.com by any other name. J. Finance 56, 2371–2388.
Deng, Zhen, Lev, Baruch, Narin, Francis, 1999. Science and technology as predictors of stock performance. Financ. Anal. J. 55, 20–32.
DeVault, Luke, Sias, Richard, Starks, Laura, 2019. Sentiment metrics and investor demand. J. Finance 74, 985–1024.
Edmans, Alex, 2011. Does the stock market fully value intangibles? Employee satisfaction and equity prices. J. Financ. Econ. 101, 621–640.
Engelberg, Joseph, McLean, R. David, Pontiff, Jeffrey, 2018. Anomalies and news. J. Finance 73, 1971–2001.
Fama, Eugene F., French, Kenneth R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116, 1–22.
Fombrun, Charles J., Shanley, Mark, 1990. What's in a name? Reputation building and corporate strategy. Acad. Manag. J. 33, 233–258.
Frankel, Richard, Litov, Lubomir, 2009. Earnings persistence. J. Account. Econ. 47, 182–190.
Gabaix, Xavier, Landier, Augustin, 2008. Why has CEO pay increased so much? Q. J. Econ. 123, 49–100.
Glushkov, Denys, 2005. Sentiment Betas. Unpublished Working Paper, University of Texas, Austin.
Gompers, Paul, Ishii, Joy, Metrick, Andrew, 2003. Corporate governance and equity prices. Q. J. Econ. 118, 107–155.
Green, Clifton T., Jame, Russell, 2013. Company name fluency, investor recognition, and firm value. J. Financ. Econ. 109, 813–834.
Harvey, Campbell R., 2017. Presidential address: The scientific outlook in financial economics. J. Finance 72, 1399–1440.
Head, Alex, Smith, Gary, Wilson, Julia, 2009. Would a stock by any other ticker smell as sweet? Q. Rev. Econ. Finance 49, 551–561.
Hirshleifer, David, Hsu, Po-Hsuan, Li, Dongmei, 2018. Innovative efficiency and stock returns. J. Financ. Econ. 107, 632–654.
Hirshleifer, David, Teoh, Siew Hong, 2003. Limited attention, information disclosure, and financial reporting. J. Account. Econ. 36, 337–386.
Hong, Harrison, Kacperczyk, Marcin, 2009. The price of sin: The effects of social norms on markets. J. Financ. Econ. 93, 15–36.
Hong, Harrison, Kostovetsky, Leonard, 2012. Red and blue investing: Values and finance. J. Financ. Econ. 103, 1–19.
Hong, Harrison, Sraer, David, 2016. Speculative betas. J. Finance 71, 2095–2144.
Hong, Harrison, Stein, Jeremy C., 2007. Disagreement and the stock market. J. Econ. Perspect. 21, 109–128.
Hsu, Ching-Yu, Huang, Chia-Wei, Lin, Chih-Yen, 2022. Company name fluency and acquisition target analysis. Eur. Financial Manag. forthcoming.
Jin, XueJun, Shen, YiFan, Yu, Bin, 2021. Stock name length and high visibility premium. Finance Res. Lett. 39, 101595.
Karpoff, Jonathan M., Rankine, Graeme, 1994. In search of a signaling effect: The wealth effects of corporate name changes. J. Bank. Financ. 18, 1027–1045.
Kaul, Aditya, Mehrotra, Vikas, Morck, Randall, 2000. Demand curves for stocks do slope down: New evidence from an index weights adjustment. J. Finance 55, 893–912.
Kile, Charles O., Phillips, Mary E., 2009. Using industry classification codes to sample high-technology firms: Analysis and recommendations. J. Account., Auditing Finance 24, 35–58.
Kumar, Alok, 2009. Who gambles in the stock market? J. Finance 64, 1889–1933.
Lev, Baruch, Sougiannis, Theodore, 1996. The capitalization, amortization, and value-relevance of R & D. J. Account. Econ. 21, 107–138.
Mueller, Holger M., Ouimet, Paige P., Simintzi, Elena, 2017. Within-firm pay inequality. Rev. Financ. Stud. 30, 3605–3635.
Muzellec, Laurent, 2006. What is in a name change? Re-joycing corporate names to create corporate brands. Corp. Reput. Rev. 8, 305–316.
Novy-Marx, Robert, 2013. The other side of value: The gross profitability premium. J. Financ. Econ. 108, 1–28.
Oppenheimer, Daniel M., 2006. Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly. Appl. Cogn. Psychol. 20, 139–156.
Pástor, Ľuboš, Stambaugh, Robert F., 2003. Liquidity risk and expected stock returns. J. Polit. Econ. 111, 642–685.
Perkins, A.G., 1995. Corporate identity: What's in a name?. Harvard Bus. Rev. 73, 11–13.
Petersen, Mitchell A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. Rev. Financ. Stud. 22, 435–480.
Rashes, Michael S., 2001. Massively confused investors making conspicuously ignorant choices (MCI–MCIC). J. Finance 56, 1911–1927.
Schwarz, Norbert, 2004. Meta-cognitive experiences in consumer judgment and decision making. J. Consum. Psychol. 14, 332–348.
Schwarz, Norbert, Jalbert, Madeline, Noah, Tom, Zhang, Lynn, 2021. Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. Consum. Psychol. Rev. 4, 4–25.
Shah, Anuj K., Oppenheimer, Daniel M., 2009. The path of least resistance: Using easy-to-access information. Curr. Direct. Psychol. Sci. 18, 232–236.
Shleifer, Andrei, 1986. Do demand curves for stocks slope down? J. Finance 41, 579–590.
Silva, Rita R., Garcia-Marques, Teresa, Mello, Joana, 2016. The differential effects of fluency due to repetition and fluency due to color contrast on judgments of truth. Psychol. Res. 80, 821–837.
Song, Hyunjin, Schwarz, Norbert, 2009. If it's difficult to pronounce, it must be risky. Psychol. Sci. 20, 135–138.
Stambaugh, Robert F., Yu, Jianfeng, Yuan, Yu, 2012. The short of it: Investor sentiment and anomalies. J. Financ. Econ. 104, 288–302.
Stambaugh, Robert F., Yuan, Yu, 2017. Mispricing factors. Rev. Financ. Stud. 30, 1270–1315.
Statman, Meir, Fisher, Kenneth L., Anginer, Deniz, 2008. Affect in a behavioral asset-pricing model. Financ. Anal. J. 64, 20–29.
Tadelis, Steven, 1999. What's in a name? Reputation as a tradable asset. Amer. Econ. Rev. 89, 548–563.
Travers, Jeffrey R., Olivier, Donald C., 1978. Pronounceability and statistical Englishness as determinants of letter identification. Am. J. Psychol. 91, 523–538.
Yu, Jianfeng, Yuan, Yu, 2011. Investor sentiment and the mean–variance relation. J. Financ. Econ. 100, 367–381.