Journal of Financial Economics 109 (2013) 813–834

Contents lists available at SciVerse ScienceDirect

Journal of Financial Economics

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

Company name fluency, investor recognition, and firm value ☆

T. Clifton Green [a,*], Russell Jame [b,c]

[a] Goizueta Business School, Emory University, 1300 Clifton Rd., Atlanta, GA 30322, USA
[b] Australian School of Business, University of New South Wales, Gate 2 High Street, Sydney NSW 2052, Australia
[c] Gatton College of Business and Economics, University of Kentucky, Lexington, KY 40506, USA

ARTICLE INFO

Article history:
Received 22 September 2011
Received in revised form
31 October 2012
Accepted 28 January 2013
Available online 12 April 2013

JEL classification:
G02
G11
G12

Keywords:
Company names
Fluency
Investor recognition

ABSTRACT

Research from psychology suggests that people evaluate fluent stimuli more favorably than similar information that is harder to process. Consistent with fluency affecting investment decisions, we find that companies with short, easy to pronounce names have higher breadth of ownership, greater share turnover, lower transaction price impacts, and higher valuation ratios. Corporate name changes increase fluency on average, and fluency-improving name changes are associated with increases in breadth of ownership, liquidity, and firm value. Name fluency also affects other investment decisions, with fluently named closed-end funds trading at smaller discounts and fluent mutual funds attracting greater fund flows.

1. Introduction

Choosing from among the thousands of stocks to invest in is a difficult decision for most people. When making complicated choices, research from psychology suggests people simplify the task by relying on mental shortcuts (Tversky and Kahneman, 1973). One input shown to be influential in the decision making process is fluency, or the ease with which people process information. Research has established that fluency has an impact on judgment that is independent of the content of the information.[1] Specifically, fluent stimuli have been shown to appear more familiar and likeable than similar but less fluent stimuli, resulting in higher judgments of preference [Alter and Oppenheimer (2009) provide a review].

The observation that fluency gives rise to feelings of familiarity and affinity suggests it could influence investor behavior. A number of studies show that investors are drawn to familiar stocks. French and Poterba (1991) find that investors overweight domestic stocks in their portfolios, and Coval and Moskowitz (1999, 2001) and Huberman (2001) find that fund managers prefer investing in locally headquartered firms.[2] In addition, evidence exists that

☆ We thank Yilin Wu for providing information on name changes and Shahryar Khan for research assistance. We also thank Michael Cooper (the referee), Gustavo Grullon, Byoung Hwang, Madhu Veeraraghavan, and seminar participants at Clemson University, Emory University, University of Miami, University of Melbourne, SAC Capital, the Behavioral Finance and Capital Markets Conference at Adelaide University, the European Finance Association meeting, and the Behavioral Science Conference at Yale University for helpful comments.

* Corresponding author. Tel.: +1 404 727 5167.
E-mail address: clifton.green@emory.edu (T.C. Green).

[1] For example, Schwarz, Bless, Strack, Klump, Rittenauer-Schatka, and Simmons (1991) ask participants to recall examples of assertive behavior and find that those asked to recall six examples (an easy task) later rate themselves as being more assertive than those asked to recall 12 examples (a difficult task). Participants emphasize ease of recall over the information gathered by the exercise.

[2] Other work that suggests familiarity can influence investment decisions includes Cooper and Kaplanis (1994), Benartzi (2001),

0304-405X/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jfineco.2013.04.007

investors prefer likeable stocks. For example, Statman, Fisher, and Anginer (2008) present a theory in which admired companies have higher valuations, and they find corresponding empirical evidence of lower returns among Fortune's most admired companies. Similarly, Hong and Kacperczyk (2009) find that sin stocks (alcohol, tobacco, and gaming companies) have lower analyst coverage and higher expected returns than otherwise comparable stocks.[3]

In this article, we investigate a new channel by which familiarity and affinity could influence investor behavior. Specifically, we examine the effects of company name fluency on breadth of ownership, liquidity, and firm value. Marketing research has long emphasized the importance of product names. For example, Bao, Shao, and Rivers (2008) find that products with easy to pronounce names exhibit increased brand recognition. Cooper, Dimitrov, and Rau (2001) suggest that the choice of company name could be important to investors as well. They find significant event period returns for firms with name changes to dotcom names during the Internet boom. In related work, Cooper, Gulen, and Rau (2005) find that mutual funds receive increased flows following name changes incorporating recently successful styles. Our emphasis is not on the information signaled by a company name but rather on the ease with which the information is processed by investors.

We hypothesize that companies with names that are easy to mentally process (i.e., fluent names) experience higher levels of breadth of ownership, improved liquidity, and higher firm values. Practically speaking, when choosing from among drug manufacturers, people could instinctively feel more comfortable investing in a name such as "Forest Laboratories" than the less fluent "Allergan Ligand Retinoid Therapeutics." We operationalize this idea by developing a measure of company name fluency based on length and ease of pronunciation. Oppenheimer (2006) finds evidence that short, simple words are processed more fluently, which activates positive affective states and influences statement evaluation. Along these lines, we reason that shorter company names are easier to process than longer names (e.g., Google versus Albuquerque Western Solar Industries), and we develop a *length* score based on the number of words in a company name.

Research in psychology suggests ease of pronunciation also has an impact on fluency and decision making. For example, Song and Schwarz (2009) ask participants to evaluate fictional food additives and amusement park rides and find that less fluent names (e.g., Hnegripitrom and Vaiveahtoishi) are considered to be riskier than more fluent choices (e.g., Magnalroxate and Chunta). In a financial setting, Alter and Oppenheimer (2006) find that survey participants predict higher future returns for fictional companies with more fluent names (e.g., Barnings versus Xagibdan).

We examine two fluency proxies that correlate with ease of pronunciation. Our first measure is the "Englishness" algorithm of Travers and Olivier (1978), which evaluates an expression based on the frequency with which its letter clusters appear in the English language. Our second approach examines whether all the words in a company name comply with a spell-check filter, based on the idea that company names that contain dictionary words are on average easier to pronounce than proper nouns or coined expressions (e.g., PharMerica or Genoptix).

We first investigate whether company name fluency affects breadth of ownership and stock liquidity. We find companies with short, easy to pronounce names have higher levels of breath of ownership, greater share turnover, and lower levels of the Amihud (2002) illiquidity measure. The results are robust to firm controls and hold among both retail investors and mutual fund managers. The results are weaker among larger firms, which is consistent with the idea that less fluent names become familiar through repeated exposure (e.g., Xerox). Together, the evidence supports the view that companies with fluent names more easily attract investors.

We next investigate the relation between fluency and firm value. We expect the familiarity and affinity associated with fluency to generate excess demand for companies with fluent names relative to companies with nonfluent names. If demand curves for stocks are downward-sloping (e.g., Shleifer, 1986; Kaul, Mehrotra, and Morck, 2000), then these differences in demand should translate into differences in valuation. Moreover, the effects of fluency on breadth of ownership and liquidity could also have important implications for firm value. For example, the Merton (1987) investor recognition hypothesis suggests breadth of ownership influences valuation. Specifically, low investor recognition leads to poor risk sharing, and the added risk leads to lower valuations and higher investment returns.[4] In other work, Amihud and Mendelson (1986) show that firms with higher levels of liquidity have lower required rates of returns and, therefore, higher firm values.

Consistent with this reasoning, firms with more fluent names have significantly higher Tobin's q and market-to-book ratios. After controlling for return on equity and other proxies for growth opportunities, we find in the cross section that a 1 unit increase in name fluency, such as reducing name length by one word, is associated with a 2.53% increase in the market-to-book ratio. For the median size company in the sample, this difference translates into \$3.75 million in added market value. Similar to the results for breadth of ownership, we find the connection between company name fluency and valuation weakens among larger firms. Moreover, we find that after controlling for breadth of ownership and liquidity, the fluency premium

(*footnote continued*)
Grinblatt and Keloharju (2001), Sarkissian and Schill (2004), Ivkovic and Weisbenner (2005), Massa and Simonov (2006), Seasholes and Zhu (2010), and Cohen (2009).

[3] There is evidence suggesting aggregate market returns are also influenced by affect. For example, Hirshleifer and Shumway (2003) find that stock market returns are higher on sunny days, and Edmans, Garcia, and Norli (2007) find that losses in soccer matches have a significant negative effect on the losing country's stock market.

[4] Several papers find empirical support for the Merton (1987) investor recognition hypothesis, including Kadlec and McConnell (1994), Chen, Noronha, and Singal (2004), and Bodnaruk and Ostberg (2009).

is cut roughly in half, which suggests that breadth of ownership and liquidity are channels through which company name fluency increases firm value.

We next investigate the effects of fluency altering name changes. The sample consists of 2,630 firms that have variation in their fluency score over time. We find that name changes significantly increase fluency on average, which is consistent with an intuitive awareness on the part of firms of the importance of name fluency. Moreover, using fixed effect regressions, we find that within-firm variation in fluency score is significantly related to breadth of ownership, liquidity, and firm value. For example, a 1 unit increase in *fluency* score is associated with a 5.80% increase in retail breadth of ownership, a 3.56% increase in total turnover, and a 0.94% increase in Tobin's q.

Our final set of tests examines whether name fluency influences other investment decisions and, in particular, the choice of investment fund. We expect that investors instinctively prefer fluently named investment funds over less fluent funds. Consistent with this conjecture, we find fluently named mutual funds receive 2.5% higher annual net inflows than less fluent funds after adjusting for past performance and other controls. Moreover, we find fluent closed-end funds trade at higher levels relative to their net asset values than less fluent funds, which provides independent evidence that name fluency affects asset prices. Consistent with the common stock results, the fluency effects on both mutual funds and closed-end funds are considerably stronger among smaller funds. Taken together, the results highlight that name fluency is an important channel by which familiarity and affinity influence investor behavior.

The remainder of the paper is organized as follows. Section 2 describes the data and our method for measuring fluency and presents descriptive statistics. Sections 3 and 4 examine the effects of company name fluency on breadth of ownership and liquidity, respectively. Section 5 examines the value implications of name fluency for stocks. Section 6 presents additional company name analysis. Section 7 explores the impact of fund name fluency on closed-end fund discounts and mutual fund flows. Section 8 concludes.

2. Data and methodology

In this section we describe the stock sample, discuss our methodology for measuring company name fluency, and present descriptive statistics.

2.1. Sample selection

The initial sample includes all securities with share-codes 10 or 11 (excluding American Depositary Receipts, closed-end funds, and Real Estate Investment Trusts) that are contained in the intersection of the Center for Research in Security Pricing (CRSP) monthly return file and the Compustat fundamentals annual file between 1982 and 2009.[5] We obtain historical company names from CRSP and begin by expanding CRSP abbreviations. For example, COMMONWEALTH TELE ENTRPS INC is changed to Commonwealth Telephone Enterprises Inc. If an abbreviation is ambiguous (e.g., TELE could stand for telephone, telecommunications, or television), we check the Securities and Exchange Commission (SEC) Electronic Data-Gathering, Analysis, and Retrieval (EDGAR) system to obtain the company legal name as reported on SEC filings. After satisfying the data requirements, the final sample consists of 14,926 companies, 18,585 unique company names, and 133,400 firm-year observations.

2.2. Measures of company name fluency

Alter and Oppenheimer (2009) define fluency as “the subjective experience of ease with which people process information” (p. 219). We are specifically interested in linguistic fluency, which concerns phonological and lexical simplicity as opposed to other forms of fluency such as visual clarity. For example, McGlone and Tofighbakhsh (2000) find that rhyming aphorisms are considered to be more true than similar non rhyming versions (e.g., *woes unite foes* versus *woes unite enemies*). Oppenheimer (2006) that finds substituting shorter and simpler alternatives for more complex words into college admission essays (e.g., *use* versus *utilize*) improves assessments of the writer's intelligence. In other work, Shah and Oppenheimer (2007) find that survey participants place more emphasis on stock recommendations from hypothetical Turkish brokerage firms with easier to pronounce names (e.g., Artan versus Lasiea).

In a similar way, we hypothesize that investors could instinctively prefer stocks with fluent company names. We measure name fluency along three dimensions. First, we reason that shorter company names are likely to be easier to mentally process. To measure company name length, we first remove expressions that are part of the legal name but are often omitted when referring to the company. Specifically, we exclude expressions such as Co., Corp., Inc., Ltd., LLC, and FSB if they are the last expression in the company name. We also exclude articles and conjunctions and articles (e.g., a, the, and, or, but), and we drop the state of incorporation, which is frequently reported in bank names. Thus, Home & City Savings Bank/NY is modified to Home City Savings Bank. We also drop hyphens between words (e.g., Wal-Mart becomes Wal Mart) and.com at the end of words. After these adjustments, we count the number of words in a company name. Company names containing one word (e.g., Google or Microsoft) are given a *length* score of 3, two words (e.g., Sun Microsystems) are given a *length* score of 2, and greater than two words (e.g., Albuquerque Western Solar Industries) are given a *length* score of 1.[6]

We also examine two measures of name fluency related to ease of pronunciation. Research in psychology and marketing typically relies on surveys to measure pronounceability. However, in our setting survey responses regarding company name pronounceability are likely to be correlated with past

[5] Prior to 1982, volume data are unavailable for Nasdaq firms. We repeat the analysis for all NYSE and Amex firms from 1963 to 2009 and find similar results.

[6] The results are very similar when using the reciprocal of the number of words in the company name to measure length.

performance and, therefore, high breadth of ownership or liquidity could lead to greater ease of pronunciation, not the other way around. We avoid this problem by relying on text-based measures of ease of pronunciation.

Our first approach is the linguistic algorithm developed by Travers and Olivier (1978) to assess the Englishness of a given word. The Englishness (E) of an n-letter string $\#L_1,L_2,\ldots,L_n\#$ (where # denotes space and L_i denotes the letter in the ith position in the string) is defined as the probability that the string will be generated by the following rule:

$$E = P(\#L_1L_2\ldots L_{n-1}L_n\#) = P(\#)P(L_1|\#)P(L_2|\#L_1)P(L_3|L_1L_2),\ldots,P(L_n|L_{n-2}L_{n-1})P(\#|L_{n-1}L_n), \quad (1)$$

where each conditional probability $P(L_k|L_{k-2}L_{k-1})$ is the probability that letter L_k follows letters L_{k-2} and L_{k-1} in printed English. Travers and Olivier (1978) and Rubin and Friendly (1986) show that Englishness is highly correlated with other measures of pronounceability and facilitates recall in tests of word recognition. Intuitively, the trigam THE appears in printed English roughly five hundred times more often than the trigram THL [i.e., $P(E|TH) > P(L|TH)$]. Thus, words that contain the trigram THE are viewed as more English than words that contain the trigam THL.

The probability expression in Eq. (1) is estimated by substituting relative bigram and trigram frequencies $F(L_{k-2}L_{k-1}L_k)/F(L_{k-2}L_{k-1})$ for $P(L_k|L_{k-2}L_{k-1})$. Negative logs are also taken to create a positive *Englishness* score (E') that generally ranges between 1 and 20. Specifically, E' is estimated as

$$E' = -\left[\log F(\#L_1L_2) + \log\frac{F(L_1L_2L_3)}{F(L_1L_2)} + \cdots + \log\frac{F(L_{n-1}L_nL\#)}{F(L_{n-1}L_n)}\right]. \quad (2)$$

We estimate $F(L_{k-2}L_{k-1}L_k)$ using data from *The Corpus of Contemporary American English,* which provides detailed estimates on the frequency of English words from over 160,000 texts from 1990 to 2010.[7] In practice, Englishness is correlated with word length, and we control for this tendency by regressing Englishness on word length and using the residuals as our measure of Englishness.

Because one highly non-English word can considerably reduce the fluency of a company name, we focus on the word with the lowest *Englishness* score within the company name. We then rank companies based on their minimum Englishness score. Companies in the bottom quintile of Englishness are given an *Englishness* score of 0, and all other companies are given an *Englishness* score of 1.

Our final measure of fluency is based on word familiarity, which is also related to ease of pronunciation. We propose that words that appear in the English dictionary are likely to be more familiar and recognizable on average than proper nouns or created expressions (e.g., PharMerica or Genoptix). To operationalize this idea, we examine whether each word within the (adjusted) company name passes through Microsoft spell-check in all lowercase letters.[8] If all words in the company name pass through the spell-check filter, then the company is given a *dictionary* score of 1. All other company names are given a *dictionary* score of 0. Our primary company name *fluency* measure is an aggregate score defined as the sum of the *length*, *Englishness*, and *dictionary* scores. Appendix A reports the five smallest and five largest companies, based on 2009 market equity, for each of the five aggregate *fluency* scores.[9]

2.3. Other variable construction

For each firm, we collect data on share price, shares outstanding, stock returns, volume, exchange membership, and Standard Industrial Classification (SIC) codes from CRSP. We obtain data on book value of equity, book value of debt, book value of assets, Standard & Poor's (S&P) 500 membership, the number of industry segments in which the firm operates, advertising expenditures, research and development (R&D) expenditures, net income, earnings before interest taxes depreciation and amortization (EBITDA), and sales from Compustat. Mutual fund breadth of ownership is computed using the Thomson Financial s12 files, and retail breadth of ownership and turnover is computed using data from a large discount brokerage (for more details, see Barber and Odean, 2000, 2001; Kumar, 2009). For each firm-year we compute a number of additional control variables. The full list of variables and the details of their construction are presented in the Appendix B.

2.4. Descriptive statistics

Table 1 presents the time series average of annual cross-sectional summary statistics computed from 1982 to 2009. In an average year, the cross section includes 4,600 firms. The average firm has a market capitalization of $1.6 billion, annual turnover of 101%, and a book-to-market ratio of 0.69. Also, the means of most of the variables are significantly larger than the medians. To reduce the effects of outliers on the analysis, we use log-transformations for most of the regression analysis.

We also present summary statistics for stocks sorted on their aggregate *fluency* score. We see that the distribution is bell-shaped with relatively few firms being either highly fluent (score=5) or highly nonfluent (score=1). In untabulated results, we find that roughly 23% of firms have a *length* score of 3, 49% of firms have a *length* score of 2, and 28% have a *length* score of 1. Roughly 34% have a *dictionary* score of 1, and by construction, 80% of firms each year have an *Englishness* score of 1. *Englishness* score and *dictionary* score are positively correlated ($\rho=0.25$), and both are negatively correlated with *length* score ($\rho=-0.07$ and -0.26, respectively).

Table 1 reveals that fluency scores also appear correlated with certain firm characteristics. Fluent companies tend to be

[7] The data set is maintained by Mark Davies, professor of corpus linguistics at Bringham Young University and is available at http://corpus.byu.edu/coca. The sample consists of the top 60 thousand English words with frequency of appearance in the corpus.

[8] We use lowercase letters to ensure that well-known company names are not recognized as words. For example, “Google” passes the spell-check filter, but “google” does not.

[9] A full list of company names and their corresponding fluency scores is available at jfe.rochester.edu/data.htm.

Table 1
Summary statistics.

The table reports the time series average of annual cross-sectional summary statistics. The sample includes all common stocks with available financial data in the Center for Research in Security Prices (CRSP) and Compustat, and it spans from 1982 to 2009. Stocks are placed into one of five groups based on their company name *fluency* score. *Fluency* scores are the sum of *length*, *Englishness*, and *dictionary* scores. Company names consisting of one, two, and more than two words receive a *length* score of 3, 2, and 1, respectively. Stocks in the bottom quintile of Englishness, as measured using a linguistic algorithm, receive an *Englishness* score of 0; all other stocks receive an *Englishness* score of 1. Company names in which all words satisfy a spell-check filter receive *dictionary* scores of 1; all other stocks receive a *dictionary* score of 0. Share price, total shares outstanding, returns, trading volume, and exchange membership are obtained from CRSP. Sales, book value of equity, earnings before interest, taxes, depreciation, and amortization (EBITDA), and total assets are obtained from Compustat. *Size* is market capitalization. *Age* is the number of months since a firm's first return appeared in the CRSP database. *Price* is share price. *Volatility* is the standard deviation of monthly stock returns over the prior year. *Turnover* is monthly volume divided by shares outstanding averaged over the previous year. *Book-to-Market* is the book value of equity divided by market capitalization. *Momentum* is the firm's equity return over the past two to 12 months. *Profitability* is EBITDA scaled by book value of assets.

	N	*Size* (millions of dollars)	*Sales* (millions of dollars)	*Age*	*Price*	*Volatility* (percent)	*Turnover* (percent)	*Book-to-Market*	*Momentum* (percent)	*Profitability* (percent)
All stocks										
Mean	4,600	1,603	1,439	153	27.91	14.11	101.00	0.69	13.30	5.12
Median		148	143	121	12.53	11.79	63.10	0.55	4.45	9.09
Standard deviation		7,962	6,525	112	643.00	9.93	159.43	1.31	63.26	27.38
Highly fluent (score=5)										
Mean	134	2,480	2,177	195	22.80	14.09	117.61	0.67	13.99	7.85
Median		254	277	182	15.36	11.89	75.36	0.53	5.18	11.44
Standard deviation		7,467	6,641	124	43.37	9.44	155.90	0.80	58.93	24.36
Fluent (score=4)										
Mean	1,590	1,742	1,577	157	17.20	14.93	114.31	0.66	13.28	4.59
Median		164	145	164	11.49	12.68	71.40	0.51	3.00	9.64
Standard deviation		9,028	7,252	113	19.49	9.87	184.34	1.40	66.60	27.21
Neutral (score=3)										
Mean	1,826	1,556	1,366	152	43.33	14.09	97.70	0.70	12.69	5.17
Median		143	146	120	12.34	11.74	61.54	0.56	4.37	9.09
Standard deviation		7,480	5,562	113	1,027.00	9.86	149.61	1.09	61.08	27.23
Nonfluent (score=2)										
Mean	898	1,380	1,273	145	18.44	13.14	86.34	0.73	14.59	5.22
Median		125	126	114	13.80	10.75	53.56	0.59	6.50	7.41
Standard deviation		6,869	6,480	107	18.67	9.83	119.61	1.25	61.64	25.21
Highly nonfluent (score =1)										
Mean	151	1,419	1,283	130	20.51	11.81	76.58	0.75	13.02	7.12
Median		148	178	105	15.79	9.62	44.33	0.62	6.69	7.73
Standard deviation		5,742	3989	108	24.52	8.72	108.85	0.82	54.91	17.74

larger, as measured by both market capitalization and sales, and older than nonfluent companies. They also tend to have higher turnover ratios, lower book-to-market ratios, and greater stock price volatility. Lastly, we see that the median fluent company tends to be more profitable than the median nonfluent company. Although the correlations are relatively modest (i.e., $\rho < 0.10$), controlling for firm characteristics in our tests is important.

3. The effects of fluency on breadth of ownership

In this section, we investigate whether investors are more likely to hold stock in companies with fluent names. Specifically, we examine whether company name fluency is related to the number of retail investors and mutual funds that own the stock. We examine this relation by estimating regressions in which the dependent variable is the natural log of the number of retail or mutual fund shareholders and the independent variables include the company name *fluency* score and other firm characteristics.[10]

Specifically, we estimate the regression specification

$$Ownership_{i,t} = a_0 + a_1 Fluency_{it-1} + \mathbf{a_2 X_{it-1}} + \varepsilon_{it},$$
$$i = 1, \ldots, N \quad t = 1, \ldots, T, \tag{3}$$

where fluency is the company's aggregate fluency score, $\mathbf{X_{it-1}}$ is a vector of firm characteristics, and ε_{it} is measurement error. Our hypothesis is that a_1 will be greater than zero. $\mathbf{X_{it-1}}$ includes a variety of firm characteristics that can help explain cross-sectional variation in breadth of ownership. For example, because breadth of ownership is likely to be strongly related to firm size, we include log (*Size*) and $[\log(Size)]^2$. Transaction costs and stock liquidity also influence the holdings of investors (e.g., Falkenstein,

[10] We examine retail and mutual fund ownership samples to investigate the effects of fluency on different investor types. An alternative

(*footnote continued*)
approach is to examine the total number of shareholders from Compustat. However, Compustat ownership data are frequently missing, and particularly for smaller firms, the effects of name fluency are likely to be stronger. In the smallest (largest) NYSE size quintile, the percentage of missing observations is 13.45% (1.83%). If we repeat the analysis using Compustat shareholder data and assume missing values are equal to five hundred (the minimum listing requirement), we find a highly significant relation between breadth of ownership and fluency score. Excluding observations with missing data produces similar but statistically weaker results.

1996), and we therefore include the reciprocal of share price (1/*Price*) and log(*Turnover*). Investors could also tilt their holdings toward value stocks, momentum stocks, older stocks, more volatile stocks, and more profitable stocks (e.g., Gompers and Metrick, 2001). Thus, we include log(*Book-to-Market Ratio*), *Momentum*, log(*Age*), log(*Volatility*), and *Profitability*. We set negative values of *Profitability* to zero and include a corresponding negative *Profitability* indicator variable. We also winsorize profitability at the 99th percentile.

Frieder and Subrahmanyam (2005) show that firms with strong brands tend to attract more shareholders; Grullon, Kanatas, and Weston (2004) show that advertising influences breadth of ownership; Kadlec and McConnell (1994) show that switching to the NYSE increases a firms' investor base; and Chen, Noronha, and Singal (2004) show that being added to the S&P 500 results in a larger investor base. To control for these effects, we include *Strong Brand*, log(*Advertising*), *NYSE*, and *S&P 500*. Because certain industries could be more visible than others, we also include dummies based on the Fama and French (1997) 49 industry classification (using two- or three-digit SIC codes produces similar results). Lastly, to control for time trends, we include year dummy variables. All variables are defined in Appendix B, and the independent variables are lagged one year relative to the dependent variable.

Table 2 presents the results of the panel regression, and T-statistics based on standard errors clustered by firm are reported in parentheses. The first column indicates a positive and significant relation between the aggregate fluency score of a company name and the number of retail shareholders. Specifically, a 1 unit increase in *fluency* score results in a 3.87% increase in the number of retail shareholders. We emphasize that the fluency coefficients reflect average benefits over the life of the firm, as company name fluency does not typically change over time.

Column 2 decomposes the *fluency* score into the *length* score, *Englishness* score, and *dictionary* score. Although the coefficient on *Englishness* is not statistically significant, both *length* score and *dictionary* score are positively and significantly related to retail breadth of ownership. Moreover, the economic magnitudes of the effects are sizable. Reducing the length of the company by one word is associated with an increase of 4.32% in retail breadth of ownership, and company names that contain all dictionary words tend to have 6.10% more shareholders than company names that contain nondictionary words.

Columns 3 and 4 repeat the analysis for mutual fund shareholders. One might expect mutual fund managers as sophisticated investors to be less prone to making investment decisions based on nonfinancial considerations. However, Coval and Moskowitz (1999) find that institutional investors prefer investing in locally headquartered firms, and Grullon, Kanatas, and Weston (2004) find that institutional investors are more likely to hold firms that advertise heavily, suggesting that sophisticated investors could also be influenced by the familiar.[11] Consistent with the retail investor results, we find that fluent companies tend to be held by more mutual fund managers. Specifically, a 1 unit increase in fluency score is associated with 2.03% increase in mutual fund breadth of ownership. The estimate is roughly half the coefficient reported for retail

Table 2
Company name fluency and breadth of ownership.

The table reports the estimates from panel regressions of the natural log of the number of retail or mutual fund shareholders on fluency and other characteristics. Retail shareholder data are obtained from a large discount brokerage data set that spans from 1991 to 1996. Mutual fund shareholder data are obtained from the CDA/Spectrum s12 database from 1982 to 2009. *Fluency* scores are the sum of *length*, *Englishness*, and *dictionary* scores. Company names consisting of one, two, and more than two words receive a *length* score of 3, 2, and 1, respectively. Stocks in the bottom quintile of Englishness, as measured using a linguistic algorithm, receive an *Englishness* score of 0; all other stocks receive an *Englishness* score of 1. Company names in which all words satisfy a spell-check filter receive *dictionary* scores of 1; all other stocks receive a *dictionary* score of 0. Detailed definitions for other control variables are presented in Appendix B. The regressions also include year dummies, a Standard & Poor's (S&P) 500 Index dummy, a NYSE exchange dummy, and industry dummies based on the Fama and French (1997) 49 industry classification. All independent variables are computed in December of the previous year. Standard errors are clustered by firm, and *t*-statistics are reported below each estimate.

Coefficients	Log(*Retail Shareholders*)		Log(*Mutual Fund Shareholders*)	
	(1)	(2)	(3)	(4)
Fluency score	3.87		2.03	
	(2.85)		(3.88)	
Length score		4.32		1.96
		(2.47)		(3.16)
Englishness score		0.27		0.93
		(0.09)		(0.67)
Dictionary score		6.10		2.94
		(2.35)		(3.28)
Log(*Size*)	−36.46	−36.26	103.69	103.70
	(−5.25)	(−5.22)	(42.97)	(42.97)
Log(*Size*)^2	3.31	3.30	−1.34	−1.34
	(10.76)	(10.73)	(−13.73)	(−13.74)
Profitability	−10.88	−10.92	39.32	39.29
	(−2.63)	(−2.65)	(8.89)	(8.88)
Log(*Turnover*)	41.49	41.46	30.23	30.22
	(31.88)	(31.88)	(48.97)	(48.91)
Log(*Book-to-Market*)	7.56	7.54	19.18	19.17
	(5.47)	(5.46)	(33.75)	(33.72)
$Momentum_{t-2,t-12}$	−3.87	−3.91	11.48	11.48
	(−4.40)	(−4.53)	(17.89)	(17.90)
Log(*Advertising*)	3.96	3.94	−1.21	−1.22
	(2.77)	(2.76)	(−2.89)	(−2.81)
Log(*Age*)	24.79	24.84	2.37	2.38
	(16.64)	(16.60)	(4.62)	(4.65)
1/*Price*	6.30	6.28	−6.53	−6.56
	(5.83)	(5.82)	(−3.09)	(−3.10)
Log(*Volatility*)	40.04	40.04	−8.13	−8.13
	(17.99)	(17.99)	(−8.64)	(−8.65)
NYSE	37.03	36.99	14.85	14.85
	(10.93)	(10.91)	(13.34)	(13.39)
S&P 500	2.20	2.20	6.42	6.61
	(0.40)	(0.40)	(4.17)	(4.26)
Strong Brand	30.96	31.15	2.59	2.61
	(3.89)	(3.91)	(1.32)	(1.33)
Industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
R^2	0.627	0.626	0.882	0.882
Clusters	6,327	6,327	11,838	11,838
Number of observations	24,289	24,289	94,549	94,549

[11] Coval and Moskowitz (1999) argue that institutional investors' preference for locally headquartered firms reflects geographical informational advantages.

shareholders, which is consistent with individual investors relying more heavily on nonfinancial criteria such as company name fluency when making investment decisions. In Column 4, we find that both *length* score and *dictionary* score are also significantly related to mutual fund breadth of ownership.

4. The effects of fluency on firm liquidity

In the previous section, we show that companies with fluent names attract a larger number of retail and mutual fund shareholders. This larger investor base could result in increased trading volume and improved liquidity. We test this hypothesis by estimating panel regressions of the natural log of either retail or total turnover on *fluency* scores and other firm characteristics as in Eq. (3). Because the decision to hold a stock and trade a stock are closely related, we use the same set of control variables as in Section 3.

The results are presented in Table 3. The first column reveals that retail turnover is significantly related to name fluency. Specifically, a 1 unit increase in fluency is associated with a 5.02% increase in retail turnover. The second column reveals that both *length* score and *dictionary* score are positive and significantly related to retail turnover. Columns 3 and 4 present the results for total turnover. Total turnover is also significantly positively related to the aggregate *fluency* score as well as all three components of fluency. A 1 unit increase in the *length* score, *Englishness* score, and *dictionary* score are associated with a 3.38%, 5.33%, and 4.32% increase in total turnover, respectively.

The significant relation between name fluency and turnover and breadth of ownership suggests that fluency could also influence analyst coverage. Consistent with this idea, we find that fluency is positively and significantly related to analyst coverage after controlling for factors known to affect coverage such as firm size, age, past returns, and exchange membership (and year and industry controls). In untabulated results, the logit regression coefficients suggest that a 1 unit change in fluency is associated with a 6.3% increase in the likelihood of analyst coverage (z-score$=3.3$). The effect of fluency on analyst coverage becomes insignificant after including mutual fund breadth of ownership and turnover as controls, which suggests it is through these channels that fluency improves analyst coverage.

The results suggest that companies with more fluent names attract more shareholders and generate greater amounts of trading. If much of this trading is unrelated to private information, then fluency could also reduce adverse selection costs, which could result in fluent stocks having smaller price impacts. To test this idea, we use the Amihud (2002) illiquidity measure as a proxy for the impact of order flow on prices.[12] Columns 5–6 report the relation between the natural log of the Amihud (2002) illiquidity measures and the *fluency* score. The results indicate that fluent firms are significantly more liquid with smaller price impacts. Specifically, a 1 unit increase in fluency reduces illiquidity by 4.61%. The illiquidity measure is also significantly negative related to the *length* score, *Englishness* score, and *dictionary* score. Taken together, the findings suggest that stocks with fluent names are more widely held and have greater levels of liquidity than similar but less fluent companies.

5. Fluency and firm value

In this section we examine the effects of company name fluency on measures of firm value. We consider several channels by which fluency may impact firm value including improvements in breadth of ownership and liquidity, as well as greater demand for the stock through increased affinity or familiarity.

5.1. Baseline specification

We investigate the effects of fluency on firm valuation by estimating regressions in which the dependent variable is a relative measure of firm value. The independent variables include the company name *fluency* score and a number of firm controls. Specifically, we estimate the panel regression

$$Value_{i,t} = a_0 + a_1 Fluency_{it-1} + \mathbf{a_2 X}_{\mathrm{it-1}} + \varepsilon_{it}, \quad i = 1, ..., N \quad t = 1, ..., T, \tag{4}$$

where fluency is the company's aggregate *fluency* score, $\mathbf{X_{it-1}}$ is a vector of firm characteristics, and ε_{it} is measurement error. Our hypothesis is that a_1 is greater than zero, which is consistent with several related hypotheses

Hypothesis 1. Fluency influences demand, and demand curves for stocks are downward-sloping (e.g., Shleifer, 1986).

Hypothesis 2. Fluency is associated with higher breadth of ownership (as shown in Section 3), and greater breadth of ownership leads to higher valuations (e.g., Merton, 1987).

Hypothesis 3. Fluency is associated with improved liquidity (as shown as Section 4), and higher liquidity results in elevated firm valuations (e.g., Amihud and Mendelson, 1986).

We consider two measures of relative value: market-to-book, which is the ratio of market value of equity to book value of equity, and Tobin's q, which is the ratio of enterprise value (debt plus market equity) to book value (debt plus book equity). We exclude observations with negative book values of equity. We take the natural log of both variables to reduce the impact of outliers.[13]

The vector of firm characteristics, $\mathbf{X_{it-1}}$, includes several variables to control for differences in growth opportunities, nontangible assets, and agency problems.[14] To

[12] Fong, Holden, and Trzcinka (2011) compare 12 low-frequency proxies that can be constructed using daily data and find that the three best proxies for price impact are the Amihud (2002) measure, the FHT Impact measure developed by Fong, Holden, and Trzcinka (2011), and the Zeros Impact measure developed by Lesmond, Ogden, and Trzcinka (1999). We find similar results when repeating the analysis using the FHT Impact and Zeros Impact measures.

[13] Hirsh and Seaks (1993) highlight that "firm and industry characteristics have multiplicative rather than additive effects on the market valuations of company assets, and provide a strong presumption for employing ln(Q) rather than Q" (p. 385). We show in Table 5 that the results are not sensitive to taking logs.

[14] The list of valuation controls is based on Edmans, Goldstein, and Jiang (2012), who also provide more detailed justifications.

Table 3
Company name fluency and liquidity.

This table reports the estimates from panel regressions of the natural log of retail turnover, total turnover, and the Amihud (2002) illiquidity measure on fluency and other characteristics. *Retail Turnover* is computed as retail share volume/shares outstanding×1000, where volume is computed from a large discount brokerage data set that spans from 1991 to 1996. *Total Turnover* is total Center for Research in Security Prices share volume/shares outstanding. The Amihud (2002) Illiquidity measure is the absolute daily return of a stock scaled by its daily total dollar volume traded, averaged across all trading days in the year. The total turnover and Amihud measure span from 1982 to 2009. *Fluency* scores are the sum of *length*, *Englishness*, and *dictionary* scores. Company names consisting of one, two, and more than two words receive a *length* score of 3, 2, and 1, respectively. Stocks in the bottom quintile of Englishness, as measured using a linguistic algorithm, receive an *Englishness* score of 0; all other stocks receive an *Englishness* score of 1. Company names in which all words satisfy a spell-check filter receive *dictionary* scores of 1; all other stocks receive a *dictionary* score of 0. Detailed definitions for other control variables are presented in Appendix B. The regressions also include year dummies, a Standard & Poor's (S&P) 500 Index dummy, a NYSE exchange dummies, and industry dummies based on the Fama and French (1997) 49 industry classification. All independent variables are computed in December of the previous year. Standard errors are clustered by firm, and *t*-statistics are reported below each estimate.

Coefficients	Log(*Retail Turnover*)		Log(*Total Turnover*)		Log(*Illiquidity*)	
	(1)	(2)	(3)	(4)	(5)	(6)
Fluency score	5.02		3.94		−4.61	
	(3.83)		(4.95)		(−4.39)	
Length score		5.65		3.38		−3.48
		(3.26)		(3.55)		(−2.75)
Englishness score		1.59		5.33		−10.39
		(0.54)		(2.48)		(−3.63)
Dictionary score		7.09		4.32		−3.44
		(2.71)		(3.07)		(−1.82)
Log(*Size*)	102.54	102.67	41.82	41.85	−90.03	−90.08
	(17.60)	(17.62)	(12.74)	(12.75)	(−15.07)	(−15.07)
Log(*Size*)^2	−4.15	−4.16	−0.66	−0.66	−2.29	−2.29
	(−16.34)	(−16.36)	(−4.58)	(−4.59)	(−8.75)	(−8.74)
Profitability	0.27	0.25	−2.46	−2.50	−20.12	−20.07
	(0.06)	(0.06)	(−1.31)	(−1.34)	(−6.71)	(−6.71)
Log(*Book-to-Market*)	−0.86	−0.87	−2.00	−1.98	−1.84	−1.88
	(−0.63)	(−0.65)	(−2.69)	(−2.67)	(−1.83)	(−1.87)
$Momentum_{t-2,t-12}$	5.18	5.15	8.72	8.73	−48.41	−48.43
	(4.81)	(4.78)	(17.36)	(17.37)	(−32.56)	(−32.56)
Log(*Advertising*)	3.85	3.84	2.33	2.34	−1.31	−1.30
	(2.73)	(2.73)	(3.26)	(3.25)	(−1.24)	(−1.23)
Log(*Age*)	3.14	3.18	−11.95	−11.92	2.06	2.01
	(2.26)	(2.29)	(−15.44)	(−15.42)	(1.98)	(1.94)
1/*Price*	−5.47	−5.48	−4.10	−4.10	5.80	5.80
	(−7.05)	(−7.11)	(−9.08)	(−9.08)	(5.88)	(5.80)
Log(*Volatility*)	79.80	79.77	66.16	66.14	−18.07	−18.10
	(37.50)	(37.47)	(65.48)	(65.43)	(−12.11)	(−12.12)
NYSE	22.47	22.41	−3.54	−3.53	−31.62	−31.62
	(6.81)	(6.79)	(−1.93)	(−1.93)	(−13.26)	(−13.27)
S&P 500	−2.30	−2.35	−1.03	−0.97	29.32	29.17
	(−0.48)	(−0.49)	(−0.37)	(−0.34)	(7.40)	(7.36)
Strong Brand	25.13	25.30	−19.28	−19.33	43.08	43.23
	(3.47)	(3.49)	(−4.99)	(−5.01)	(7.08)	(7.10)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.269	0.269	0.434	0.434	0.850	0.850
Clusters	6,860	6,860	14,044	14,044	14,037	14,037
Number of observations	26,412	26,412	115,341	115,341	115,269	115,269

control for growth opportunities, we include *Growth*, defined as sales growth over the past three years, log(*Age*), and log(*Sales*). We also include a firm's *Profitability* (EBITDA/assets). We set negative values of *Profitability* to zero and include a corresponding negative *Profitability* indicator variable. We also winsorize *Profitability* at the 99th percentile. Firms with high R&D could also have better growth options. Moreover, R&D is an intangible asset that is often not captured in the book value. Similarly, advertising and strong product brands could increase firm value through improved recognition but do not have a direct effect on book value. Lastly, firms with high asset turnovers likely have a large amount of intangible assets, which is probably to be associated with a low book value and a high Tobin's *q*. To control for these effects, we include *R&D/Sales*, *Advertising/Sales* (both winsorized at the 99%), *Strong Brand*, and *Asset Turnover* (winsorized at the 1st and 99th percentile).

To control for agency problems, we include *Leverage* and *Payout*. Both reduce free cash flows available to the manager and, therefore, limit the manager's ability to implement value destroying investment decisions. *Leverage* and *Payout* are both winsorized at the 1st and 99th percentile. We control for the diversification discount (e.g., Lang and Stulz, 1994) by including the log of the total number of industry segments in which the firm operates.

We also include *NYSE* and *S&P 500* because exchange membership and index membership could affect a firm's investor base and liquidity. Lastly, we include year dummies and industry dummies based on the Fama and French (1997) 49 industry classification. All independent variables are lagged one year relative to the dependent variable.

Table 4 presents the results of the panel regression, and *t*-statistics based on standard errors clustered by firm are reported in parentheses. The first column indicates that Tobin's *q* is positive and significantly related to *fluency* scores. A 1 unit increase in *fluency* score is associated with a 1.90% increase in Tobin's *q*. Moreover, all three components of the *fluency* score are significantly and positively related to Tobin's *q*. Not surprisingly, Columns 3 and 4 reveal a similar relation between *fluency* score and *market-to-book ratio*. A 1 unit increase in *fluency* score is associated with a 2.53% increase in the *market-to-book ratio*. For the median size company in the sample this difference translates into $3.75 million in added market capitalization.

Table 4
Company name fluency and firm value.

The table reports the estimates of panel regressions of the natural log of *Tobin's q* or *Market-to-Book* on fluency and other characteristics. *Tobin's q* is the ratio of the enterprise value (market value of equity plus debt) to book value (debt plus book equity). *Market-to-Book* is the market value of equity divided by book value of equity. *Fluency* scores are the sum of *length*, *Englishness*, and *dictionary* scores. Company names consisting of one, two, and more than two words receive a *length* score of 3, 2, and 1, respectively. Stocks in the bottom quintile of Englishness, as measured using a linguistic algorithm, receive an *Englishness* score of 0; all other stocks receive an *Englishness* score of 1. Company names in which all words satisfy a spell-check filter receive *dictionary* scores of 1; all other stocks receive a *dictionary* score of 0. Detailed definitions for other control variables are presented in Appendix B. The regressions also include year dummies, a Standard & Poor's (S&P) 500 Index dummy, a NYSE exchange dummy, and industry dummies based on the Fama and French (1997) 49 industry classification. All independent variables are computed in December of the previous year. Standard errors are clustered by firm, and *t*-statistics are reported below each estimate.

Coefficients	Log(*Tobin's q*)		Log(*Market-to-Book*)	
	(1)	(2)	(4)	(5)
Fluency score	1.90		2.53	
	(5.19)		(4.12)	
Length score		2.17		2.72
		(4.55)		(3.45)
Englishness score		1.57		2.91
		(1.87)		(2.18)
Dictionary score		1.69		1.80
		(2.32)		(1.49)
Log(*Sales*)	−5.33	−5.33	−9.72	−9.73
	(−19.93)	(−19.97)	(−22.69)	(−22.69)
Profitability	3.01	3.01	4.71	4.71
	(61.80)	(61.81)	(66.03)	(66.02)
Log(*Age*)	−4.56	−4.58	−6.85	−6.88
	(−9.84)	(−9.87)	(−8.93)	(−8.94)
Sales Growth	1.24	1.23	3.22	3.22
	(2.69)	(2.68)	(2.39)	(2.38)
Asset Turnover	−1.87	−1.86	−1.04	−1.03
	(−3.62)	(−3.61)	(−1.17)	(−1.16)
R&D/Sales	2.30	2.29	2.21	2.20
	(4.11)	(4.09)	(3.31)	(3.30)
Advertising/Sales	0.56	0.55	0.99	0.99
	(3.68)	(3.67)	(4.47)	(4.46)
Log(*Segments*)	−1.23	−1.24	−2.70	−2.72
	(−2.81)	(−2.86)	(−3.50)	(−3.53)
Leverage	5.72	5.73	106.34	106.37
	(3.33)	(3.34)	(36.74)	(36.77)
Payout	−0.99	−0.99	−0.04	−0.04
	(−2.28)	(−2.28)	(−0.04)	(−0.05)
NYSE	5.13	5.12	12.19	12.19
	(6.30)	(6.30)	(8.34)	(8.34)
S&P 500	17.58	17.58	32.13	32.12
	(17.55)	(17.53)	(17.62)	(17.62)
Strong Brand	19.38	19.42	35.79	35.79
	(9.81)	(9.83)	(11.35)	(11.35)
Industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
R^2	0.345	0.345	0.289	0.289
Clusters	13,422	13,422	13,422	13,422
Number of observations	110,491	110,491	110,491	110,491

5.2. Alternative specifications

In Table 5, we examine the robustness of the relation between *fluency* score and firm value. For the sake of brevity, in each row we now report only the coefficient estimate on *fluency* score and any new variables added to the specification. We report results for Tobin's *q*. The results for market-to-book are very similar.[15] For reference, the first row of Table 5 reports the coefficient and *t*-statistic on *fluency* score from the baseline specification.

In Row 2, we repeat the analysis using the Fama and Macbeth (1973) methodology. Specifically, we estimate cross-sectional regressions each year and average coefficients across years. We use the Newey and West (1987) adjustment for serial correlation with the maximum possible lag-length. The estimate from the Fama and Macbeth regression is similar in magnitude to the panel regression result and is highly significant. Also, the standard error from the Fama and MacBeth estimate is significantly smaller than the standard error from the baseline specification, which highlights the importance of computing standard errors clustered by firm.

Table 1 reveals that *fluency* score is correlated with sales, age, and profitability. If a nonlinear relation exists between these control variables and Tobin's *q*, then the coefficient on fluency in the linear specification could incorrectly reflect the influence of these other characteristics. To address this concern, we create a size, age, and profitability (SAP)–adjusted measure of Tobin's *q*. Specifically, at the end of each year, we sort all stocks into one of five size quintiles based on NYSE sales. Within each size quintile, we divide all stocks into quintiles based on age. Finally, within each of the 25 size and age portfolios, firms are sorted into quintiles based on profitability, yielding 125 portfolios. The benchmark for each company is the portfolio to which it belongs. The SAP-adjusted Tobin's *q* for each firm is the difference between the firm's Tobin's *q* and the equally weighted average Tobin's *q* of its benchmark portfolio. We then repeat the panel regression, in which the dependent variable is SAP–adjusted Tobin's *q*. The results, presented in Row 3, indicate that this adjustment has virtually no impact on the *fluency* score coefficient.

[15] We also repeat this type of analysis for breadth of ownership and liquidity. The results in Tables 2 and 3 are robust to various specifications.

Table 5
Company name fluency and firm value: robustness checks.

This table presents the results of variations on the pooled regression in Table 4. The dependent variable is the natural log of Tobin's q (unless stated otherwise). Row 1 reports the results from the main specification reported in Table 4. Row 2 reports the coefficients using the Fama and MacBeth (1973) methodology. Row 3 reports results in which the dependent variable is adjusted by subtracting the mean Tobin's q from one of 125 benchmark portfolios matched on size, age, and profitability. Row 4 repeats the baseline specification but excludes financials [Standard Industrial Classification (SIC) of 6000–6999], and Row 5 removes firms with headquarters outside the US. Row 6 winsorizes the dependent variable at the 1st and 99th percentile. Row 7 uses Tobin's q (not in logs) as the dependent variable. Rows 8 and 9 employ finer industry control partitions (dummy variables based on three- and four-digit SIC codes). Rows 10 through 12 add log of turnover and log of mutual fund shareholders. With the exception of Row 2, t-statistics, based on standard errors clustered by firm, are reported in parentheses. In Row 2, t-statistics are computed from the time series standard deviation of annual coefficient estimates with a Newey and West (1987) adjustment for serial correlation.

Row	Specification	*Fluency* score	Turnover	Mutual fund shareholders
1	Baseline specification	1.90		
		(5.19)		
2	Fama and MacBeth estimates	1.64		
		(5.14)		
3	Size, age, and profitability—adjusted Tobin's q	1.89		
		(5.25)		
4	Remove financials	2.04		
		(4.77)		
5	Remove foreign firms	1.88		
		(5.12)		
6	Winsorize q	1.82		
		(5.10)		
7	Raw q (not in logs) ×100	4.88		
		(3.49)		
8	Include three-digit SIC dummies	1.49		
		(4.22)		
9	Include four-digit SIC dummies	1.30		
		(3.60)		
10	Include total turnover	1.40	10.08	
		(3.96)	(31.17)	
11	Include breadth of ownership	1.31		13.68
		(3.85)		(35.03)
12	Include turnover and breadth of ownership	1.12	5.99	11.28
		(3.30)	(18.00)	(28.10)

Our next robustness check involves removing financials. Because the meaning of certain control variables are often different for financial companies (e.g., leverage), in Row 4, we repeat the analysis excluding all financial companies (SIC code 6000–6999). The *fluency* score coefficient increases slightly, indicating that the results are not driven by financial firms. An additional concern is that foreign firms could be more likely to bear nonfluent names, in which case the results could be related to home bias (e.g., French and Poterba, 1991). In Row 5, we exclude all firms with headquarters outside the United States and find the coefficient on *fluency* score remains essentially unchanged.

To verify that the results are not driven by outliers, in Row 6 we winsorize the log of Tobin's q at the 1% and 99% percentile. The coefficient on *fluency* score remains very similar, suggesting that the results are not driven by outliers. Row 7 repeats the analysis using the raw value (i.e., the nonlogged value) of Tobin's q. A 1 unit increase in *fluency* score is associated with a 0.049 increase in Tobin's q. The average (median) firm has a Tobin's q of 2.06 (1.30). Thus, a 0.049 in Tobin's q corresponds to a 2.38% (3.77%) increase, both of which are larger than the 1.90% predicted increase in the baseline specification (using a winsorized value of q leads to similar conclusions).

In Rows 8 and 9, we add three- and four-digit SIC dummies and limit the analysis to industries with at least three firms. Adding finer industry partitions does reduce the coefficient on fluency, although this effect is not surprising. Using four-digit SIC codes results in 340 different industry dummies with the average (median) industry containing 11 (5.5) different firms per year. If the median industry contains only 5.5 firms, then much of the variation in *fluency* scores occurs at the industry level instead of within, which is (unfairly) captured by the industry dummies. Despite this high hurdle, the coefficient on *fluency* score remains highly significant, suggesting that, even within finely partitioned industries, a significant relation exists between company name fluency and firm value.[16]

An additional concern is that SIC codes might not fully capture the relatedness of firms in the product market space. We address this concern using the text-based network industry classification (TNIC3) developed in Hoberg and Phillips (2010a, 2010b).[17] The industry classification is based on a web-crawling and text-parsing algorithm that processes the text in the business descriptions of 10-K

[16] Another type of industry analysis is to estimate the effects of fluency separately by industry. We run separate panel regressions for each of the 49 Fama and French industries and find that the coefficient on *fluency* score is positive in 36 (73%) of the regressions.

[17] The Hoberg Phillips data library is available at http://www.rhsmith.umd.edu/industrydata.

annual filings from 1996 to 2008. During the 1996–2008 subperiod, the coefficient (t-statistic) on *fluency* score when including TNIC3 dummies is 1.66 (3.59). Over the same period, the coefficient (t-statistic) on *fluency* score when including three-digit SIC dummies, which have the same level of industry coarseness as the TNIC3 classification, is 1.98 (4.25). Thus, a more precise control for product-market relatedness does reduce the coefficient on *fluency* score by roughly 15%, yet the effects of fluency on firm value remain highly significant.

In Row 10, we include turnover. If liquid firms have higher valuations and fluency is related to higher liquidity (Hypothesis 3), then the coefficient on turnover should be positive and the coefficient on fluency should decline in magnitude. Consistent with this prediction, we find that turnover is strongly related to firm value, and the coefficient on *fluency* score falls from 1.90 to 1.40. In Row 11, we include mutual fund breadth of ownership. If breadth of ownership is positively related to firm value and fluency is related to breadth of ownership (Hypothesis 2), then the coefficient on breadth of ownership should be positive and the coefficient on fluency should be reduced. The findings from Row 11 are consistent with this prediction. Lastly, in Row 12, we include both turnover and breadth of ownership together. Both turnover and breadth of ownership remain highly significant, and the coefficient on fluency drops to 1.12. The results suggest that breadth of ownership and liquidity are two channels through which the fluency of a company name influences firm value. However, the coefficient on *fluency* score remains economically and statistically significant, which suggests company name fluency could increase firm value over and above its influence on breadth of ownership and liquidity.

5.3. *Implications for expected returns*

The impact of name fluency on firm valuation raises the question of whether it influences stock returns as well. Consider a company with a *fluency* score of 1 that generates earnings of \$1 a year in perpetuity and is priced at \$20. This corresponds to a discount rate of 5%. Now consider a company with a *fluency* score of 5 that also generates earnings of \$1 a year in perpetuity. The market-to-book estimates suggest that the fluent company should trade at a 10.12% premium (2.53×4), implying a price of \$22.02 and a corresponding discount rate of 4.54%. The difference in returns of 46 basis points per year (or roughly 1 basis point per month per unit change in *fluency* score) is unfortunately too small to easily detect statistically given the observed variation in returns. Nevertheless, we investigate the relation between company name fluency and returns empirically with Fama and MacBeth (1973) regressions each year from 1982 to 2009 of monthly returns on *fluency* score and find no significant relation between fluency and returns.

6. Additional company name analyses

In this section we provide additional evidence regarding the effects of name fluency on investor recognition and firm value. In Section 6.1 we examine whether fluency has a non-linear effect on our variables of interest. In Section 6.2 we examine whether the effects of fluency vary with firm size, and in Section 6.3 we present an analysis of fluency-altering name changes.

6.1. *Fluency sorts*

Thus far, the regressions have assumed a linear relation between measures of fluency and ownership, liquidity, and firm value. Although research from psychology suggests investors could instinctively prefer companies with more fluent names and avoid less fluent names, it is not clear that the relation should be symmetric. Some experiments emphasize the disadvantages of nonfluent stimuli (e.g., higher perceived risk; see Song and Schwarz, 2009), whereas others focus on the benefits of fluency (e.g., more likely to be perceived as correct; see McGlone and Tofighbakhsh, 2000).[18]

In this subsection, we explore whether the relation between name fluency and breadth of ownership, improved liquidity, and valuation ratios is driven primarily by investors' preference for fluent stocks or their aversion to nonfluent socks. In particular, we examine abnormal breadth of ownership, abnormal liquidity, and abnormal firm value for portfolios of stocks sorted on *fluency* score. Abnormal breadth of ownership is defined as the observed breath of ownership less the predicted breadth of ownership, in which predicted breadth of ownership is estimated from the regression in Table 2 but excluding *fluency* score as an independent variable. Similarly, abnormal liquidity and abnormal firm value are the residuals from the regression models outlined in Tables 3 and 4, where *fluency* score is again omitted.

Fig. 1 plots the results of the analysis. For ease of interpretation, we multiply the coefficient on the Amihud (2002) illiquidity measure by negative one. The results reveal a generally monotonic relation for each dependent variable. However, fluency has an asymmetric effect on breadth of ownership and liquidity. For example, mutual fund breadth of ownership is 1.66% higher than expected in highly fluent stocks, but 7.84% lower than expected in highly nonfluent stocks. Similarly, abnormal turnover is 5.15% in highly fluent stocks, but −14.40% in highly nonfluent stocks. A similar asymmetric pattern can be found amongst retail breadth of ownership and retail turnover. This suggests that investors are particularly repelled by highly nonfluent company names.[19]

The asymmetric effect of fluency is more muted for the valuation ratios. The abnormal Tobin's q for highly fluent (highly nonfluent) stocks is 2.42% (−3.37%). The lack of strong negative valuation effects for nonfluent companies is not surprising, given that retail investors do not typically

[18] Alter and Oppenheimer (2009) provide a review of the fluency literature.

[19] The strong response to highly nonfluent names raises the concern that the results could be driven by a relatively small number of observations. We also repeat the analysis using a three-point *fluency* score by collapsing the highest (lowest) two fluency groups into one high (low) group. Using the full set of controls, we find the following coefficients (with t-statistics in parentheses): *Retail Breadth*, 4.03 (2.65); *Mutual Fund Breadth* 2.19 (3.78); *Retail Turnover*, 5.04 (3.33); *Total Turnover*, 4.12 (4.70); Amihud's *Illiquidity*, −4.39 (−3.77); *Tobin's q*, 2.17 (5.05); and market-to-book, 2.84 (4.04).

short stocks, and their aversion to highly non-fluent stocks could be offset through purchases from sophisticated investors. Taken together, the results suggest the effects of fluency arise both from a preference for fluent company names and a somewhat stronger aversion to nonfluent names.

6.2. Company name fluency and firm size

In this subsection, we explore whether the effects of name fluency on investor recognition and firm value vary with firm size. Research from psychology indicates that previous exposure to concepts increases their fluency. For example, Labroo, Dhar, and Schwarz (2008) find priming survey participants with the concept of a frog led them to process a wine bottle with a frog on its label more favorably. In our context, we expect repeated exposure to the names of more visible companies increases the fluency of their perhaps otherwise nonfluent names (e.g., Xerox). We, therefore, expect a stronger relation between the name fluency measures and investor recognition among small firms.

In addition to being less visible, small stocks tend to have a greater concentration of retail ownership. Individual investors have been shown to be more susceptible to cognitive biases than institutional investors (e.g., Battalio and Mendenhall, 2005; Grinblatt and Keloharju, 2001), and they could rely more heavily on nonfinancial criteria such as company name fluency when making investment decisions. Finally, limits to arbitrage are also more severe for small stocks and, therefore, name fluency could have a larger influence on firm value among small firms.

We operationalize this idea by repeating the previous analysis using separate regressions on microcaps, small firms, and large firms. Following Fama and French (2008), we define microcaps as stocks with market caps below the 20th NYSE percentile. Small stocks are those with market caps between the 20th and 50th percentiles, and large stocks are those above the NYSE median.

The results of this analysis are reported in Table 6. The regressions include the full list of independent variables used as controls in Tables 2–4, but for brevity we report only the coefficient on *fluency* score. We also report *t*-statistics and the total number of observations. The results present strong evidence that the impact of fluency is considerably stronger for smaller stocks. For six of the seven variables, the effect of fluency is strongest among microcap stocks. Similarly, for five of the seven variables, the effect is weakest among large stocks. For example, a 1 unit increase in fluency increases market-to-book by 2.81% for microcap stocks, 1.10% for small stocks, and 0.87% for larger stocks.[20]

The strong effect of name fluency on microcap stocks is consistent with our predictions, yet it raises concerns about the economic significance of the results. While roughly 60% of all stocks in the sample are microcaps, they account for only 3% of the total market cap of all stocks in the sample. As an additional test, we repeat the analysis excluding microcap stocks. After excluding the 60% of the sample, in which the effects are the strongest, we continue to find a strong relation between fluency and the variables of interest. The relation is in the predicted direction for all seven variables and is statistically significant (at a 10% level) for six of the seven variables.

Although the fluency effects are stronger in microcap stocks in percentage terms, the level effects are typically larger outside of microcaps. For example, the median size for a microcap firm (over the entire sample period) is \$52 million as compared with \$890 million for non-microcap stocks. Thus, a 1 unit increase in fluency translates into an additional \$1.5 million ($\$52 \times 2.81$) in market equity for microcap stocks compared with \$10 million ($\$890 \times 1.22$) for non-microcap stocks. Thus, while the fluency coefficients are largest among microcap stocks, the evidence suggests name fluency is economically relevant for larger stocks as well.

6.3. Name changes

An alternative approach to examine the impact of fluency on breadth of ownership, liquidity, and firm value is to examine companies that have changed their name. By focusing exclusively on within-firm variation, we can address the concern that companies with fluent names are systematically different from companies with nonfluent names. Despite their conceptual appeal, in practice name changes are rarely exogenous. For example, name changes could be motivated by corporate events such as mergers or the desire to communicate a shift in business focus to market participants. It is worthwhile to examine whether fluency-enhancing mergers lead to greater investor recognition than fluency-reducing mergers, yet endogenous name changes such as these challenge our assumption that firm fundamentals do not change around the event. Therefore, in the analysis we also identify a subset of name changes that are unlikely to be related to fundamental shifts in business operations.

We begin by classifying the set of fluency-altering name changes from 1980 to 2008 into four categories: Corporate Restructure, Broad Focus, Narrow Focus, and Rebranding. Corporate Restructure name changes are driven by corporate events such as mergers (e.g., from AOL to AOL Time Warner), corporate restructurings, and other confounding events such as changes in legal status. Broad Focus name changes are name changes that are motivated by the company expanding their business lines. For example, Apple Computer changed its name to Apple to emphasize that it was expanding beyond

[20] We also use firm age as a measure of visibility and find similar but slightly weaker results. The fact that fluency has a larger effect on valuation for younger firms is consistent with Alter and Oppenheimer (2006), who find larger first day returns for fluently named Initial Public Offerings (IPOs). They examine a relatively small sample (89 observations) and rely on surveys to gauge name fluency. However, name recognition in surveys could be influenced by firm performance, and their methodology does not include many controls common in the IPO literature. We examine the relation between first-day returns and name fluency following the methodology of Green and Hwang (2012). We find a positive, but statistically insignificant, relation between IPO first-day returns and aggregate fluency score.

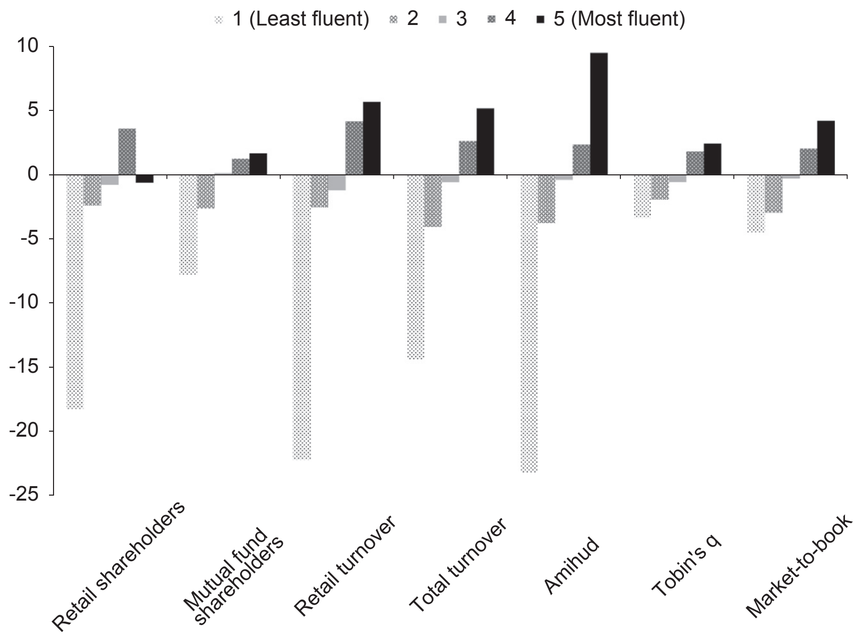

Fig. 1. The effects of company name fluency on breadth of ownership, liquidity, and firm value: portfolio sorts. This figure plots average abnormal breadth of ownership, abnormal liquidity, and abnormal firm value for portfolios sorted on *fluency* score. Abnormal breadth of ownership, abnormal liquidity, and abnormal firm value are defined as the residuals from the regressions in Tables 2–4 excluding *fluency* score as an independent variable. For ease of interpretation, the estimates for the Amihud (2002) illiquidity measure have been multiplied by negative one.

Table 6
The effects of fluency by firm size.

The table reports the estimates of panel regressions of breadth of ownership, liquidity, and valuation on fluency and other firm characteristics by firm size. The breadth of ownership, liquidity, and valuation regressions are run as specified in Tables 2–4 but are now run on subsets of stocks based on a stock's market capitalization. Panel A reports the results for microcap stocks. Panels B and C report the results for small stocks and large stocks, respectively. Panel D presents the results for all stocks excluding microcap stocks (i.e., small and large stocks). We define microcaps as stocks with market cap below the 20th NYSE percentile. Small stocks are those with market caps between the 20th and 50th percentile, and large stocks are those above the NYSE median. Standard errors are clustered by firm, and *t*-statistics are reported below each estimate in parentheses. The number of observations also is reported.

	Retail shareholders	Mutual fund shareholders	Retail turnover	Total turnover	Amihud's illiquidity	Tobin's *q*	Market-to-book
Panel A: Microcap stocks							
Fluency score	5.71	1.00	6.38	4.40	−4.34	2.22	2.81
	(3.54)	(1.51)	(3.61)	(4.38)	(−3.46)	(4.97)	(3.75)
Number of observations	14,022	50,966	15,293	68,885	68,813	65,352	65,352
Panel B: Small stocks							
Fluency score	2.53	2.41	4.40	2.26	−2.68	0.75	1.10
	(1.06)	(2.66)	(2.00)	(1.56)	(−1.47)	(1.77)	(1.54)
Number of observations	5,441	21,792	5,593	23,599	23,599	22,767	22,767
Panel C: Large stocks							
Fluency score	0.04	2.06	1.05	1.60	−1.54	0.79	0.87
	(0.02)	(3.08)	(0.51)	(1.35)	(−1.15)	(1.46)	(0.93)
Number of observations	4,841	21,785	4,917	2,957	22,957	22,504	22,504
Panel D: All but microcap stocks							
Fluency Score	1.76	2.27	3.16	2.26	−2.43	0.92	1.22
	(0.94)	(3.68)	(1.94)	(2.18)	(−1.92)	(2.28)	(1.78)
Number of observations	10,282	43,577	10,510	46,556	46,556	45,271	45,271

the computer industry. Similarly, Narrow Focus name changes are motivated by the company reducing one or more existing business lines, such as when Epix Medical changed its name to Epix Pharmaceutical to emphasize its increasing focus on developing pharmaceutical products. We consider that Corporate Restructure name changes, as well as Broad Focus and Narrow Focus name changes, could lead to fundamental shifts in business operations. For example, Wu (2010) finds that Narrow Focus name change firms have a higher Tobin's q on average after they refocus their business operations.

We classify the remaining name changes, which are less likely to be influenced by fundamental shifts in business operations, as Rebranding. Examples include adopting a recognizable brand name as the company name (e.g., from Federated Department Stores to Macy's), modifying an existing name to a shorter, simpler version (e.g., from Kaufman and Broad Home Corporation to KB Home), and completely changing the existing name while maintaining the same business model (e.g., from Quotesmith.com to Insure.com).

We use the name change classification data from Wu (2010) for the period 1980–2000, graciously provided by Yilin Wu, and extend the data through 2009 using Dow Jones Newswire searches near the event. Together, we are able to classify 2,630 fluency-altering name changes, of which 52% are Corporate Restructure, 16% are Broad Focus, 8% are Narrow Focus, and the remaining 23% are Rebranding. Across all name changes, the average fluency of a company name increases by a statistically significant 0.14 (t-statistic = 5.01). Fluency increases dramatically for Broad Focus name changes and decreases significantly for Narrow Focus name change. This result is intuitive, as Broad Focus name changes often shorten the company name (and become more fluent) by removing a word to become more general (e.g., from Candela Laser Corporation to Candela Corporation) and Narrow Focus name changes often add a word to become more specific (e.g., from Vaughn Incorporated to Vaughn Communications Incorporated).

We find that Rebranding name changes significantly enhance average fluency by 0.18 (t-statistic = 3.26). Moreover, Chief Executive Officers (CEOs) often mention fluency-related concepts in motivating Rebranding name changes, which is consistent with the idea that they seek to improve visibility instead of signaling a shift in business operations. For example, in motivating the name change from International Remote Imaging Systems to IRIS International, the CEO of IRIS International stated: "We believe that 'IRIS International' has a ring of familiarity and also reflects our growing international presence. It will further strengthen our brand and recognition among our customers and within the investment community."

We next examine how within-firm variation in *fluency* score effects breadth of ownership, liquidity, and valuation. Specifically, we estimate the effects of fluency on breadth of ownership using the regression

$$Ownership_{i,t} = a_i + a_1 Fluency_{it} + \mathbf{a_2 X_{it}} + \varepsilon_{it}, \quad i = 1, \ldots, N \quad t = 1, \ldots, T. \tag{5}$$

In contrast to the regressions in Table 2, such as Eq. (3), we now allow firm-specific intercepts (i.e., firm fixed effects), and thus our coefficients are based entirely on within-firm variation over time. $\mathbf{X_{it}}$ is a vector of firm characteristics that include all the control variables used in Table 2, but we also add dummy variables that indicate the type of name change (e.g., Rebranding, Corporate Restructure, Broad Focus, and Narrow Focus). Lastly, unlike Table 2, in which the independent variables are lagged one year, in Table 7 all the independent variables are contemporaneous to the dependent variables. This change is made to ensure that the results capture any effects that occur in the year in which the name change took place.

Table 7 reports the results of the analysis. We run the fixed effect regressions for our full sample of name changes as well as for each type of name change. The regressions use all of the independent variables, including controls for changes in fundamentals such as *Sales*, *Profitability*, and *Advertising*, although for brevity the table reports only the coefficients on *fluency* score. The central finding is that changes in *fluency* score are positively and significantly related to changes in breadth of ownership, liquidity, and firm value. For example, over the full sample of name changes, we find a significant relation between name fluency and all of the dependent variables. Moreover, the economic magnitudes are generally similar to the between-firm estimates from the panel regressions. For the subset of Rebranding name changes, the coefficient has the correct sign for all seven dependent variables and is statistically significant in five out of seven cases.[21] The results for Corporate, Broad Focus, and Narrow Focus name changes typically point in the right direction, although fewer estimates are reliably different from zero.

Taken together, the results from the name change analysis help alleviate concerns that the between-firm estimates are driven by omitted variables and confirm the importance of name fluency. In particular, the findings suggest that name fluency is an important consideration when attempting to improve visibility through a name change.

7. Investment fund name fluency and investor recognition

In this section, we explore the effects of name fluency on investor recognition and asset values in an alternative setting. Specifically, we examine whether investors exhibit a preference for fluently named investment funds, leading to larger premiums for fluently named closed-end funds and more flows for fluent open-end mutual funds.

Closed-end fund premiums offer important advantages over stock ratio analysis for examining whether name fluency affects valuation. Closed-end fund book values (net asset values) are based on the market prices of the underlying securities held by the fund. As a result, closed-end fund premiums generally are less influenced by accounting conventions, non-tangible assets (e.g., a firm's brand recognition), or differences in growth opportunities than market-to-book ratios. Closed-end funds also tend to be held primarily by retail investors who are likely more prone to making

[21] We repeat the analysis excluding observations in which the company changed its name to a strong product brand (i.e., a brand name on the Interbrand or Brandirectory list of top global brands) and find very similar results.

Table 7
The effects of fluency-altering company name changes on liquidity and firm value.

The table reports the estimates of fixed effect panel regressions of breadth of ownership, liquidity, and valuation on fluency and other firm characteristics. Fluency-altering name changes from 1980 to 2008 are classified into four categories by reading newswire descriptions of the event. Corporate Related name changes arise from corporate events such as mergers, acquisitions, and other corporate restructuring. Broad (Narrow) Focus name changes are the result of a company's decision to expand (narrow) the scope of its business lines. Name changes that do not involve fundamental shifts in business operations are classified as Rebranding. The breadth of ownership, liquidity, and valuation regressions are run as specified in Tables 2–4, with the addition of dummy variables for each firm and a dummy variable that captures the type of name change. The *fluency* score coefficients thus measure the effects of fluency altering name changes. This table reports the fluency coefficients and *t*-statistics for the full sample of name changes (Panel A), as well as the coefficient for each specific type of name change (Panels B–E). For brevity, coefficients on all other control variables are not reported. For each regression, the number of unique name changes with nonmissing data also is reported.

	Retail shareholders	Mutual fund shareholders	Retail turnover	Total turnover	Amihud's illiquidity	Tobin's *q*	Market-to-book
Panel A: All name changes							
Fluency score	5.80	1.46	6.67	3.56	−5.10	0.94	1.26
	(3.01)	(2.57)	(2.45)	(5.87)	(−5.68)	(2.48)	(2.10)
Name change observations	861	2,114	861	2,364	2,364	2,356	2,356
Panel B: Rebranding name changes							
Fluency Score	11.38	1.41	11.56	3.85	−3.85	2.73	2.65
	(2.74)	(1.28)	(1.92)	(3.25)	(−2.20)	(3.50)	(2.12)
Name change observations	162	434	162	476	476	476	476
Panel C: Corporate name changes							
Fluency Score	5.28	1.05	−3.48	4.37	−7.21	1.33	0.99
	(1.79)	(1.24)	(−0.81)	(4.80)	(−5.34)	(2.21)	(1.04)
Name change observations	527	1,201	527	1,363	1,363	1,355	1,355
Panel D: Broad focus name changes							
Fluency Score	−0.49	−1.32	16.28	3.87	−8.81	−0.94	1.24
	(−0.07)	(−1.01)	(1.60)	(2.74)	(−4.19)	(−0.91)	(0.79)
Name change observations	130	325	130	354	354	354	354
Panel E: Narrow focus name changes							
Fluency Score	−10.51	0.35	−24.16	1.55	2.15	2.86	0.03
	(−1.01)	(0.19)	(−1.65)	(0.78)	(0.72)	(1.95)	(0.01)
Name change observations	42	154	42	171	171	171	171

investment decisions based on non-fundamental information (e.g., Lee, Shleifer, and Thaler, 1991).

Open-end mutual funds provide an additional test for whether name fluency affects investor recognition. If investors instinctively avoid less fluent mutual fund names or are drawn to fluent names, we would expect to see greater fund flows into fluently named mutual funds after conditioning on performance and other controls.

7.1. Measuring fund name fluency

We obtain data on closed-end funds from 1994 [first year Compustat reports net asset values (NAVs)] to 2009 from Morningstar and Compustat. For each closed-end fund, Morningstar provides both a family name (e.g., BlackRock and ALPS Advisors) and a corresponding fund name (e.g., BlackRock High-Income and Liberty All-Star Growth). Because both fund name and family name fluency potentially influence investor decisions, we limit the sample to funds for which the family name is also present in the fund name.

We also collect data on open-end mutual fund names from 1992 [first year CRSP reports monthly total net assets (TNAs)] to 2009 from the CRSP Mutual Fund Database. To facilitate comparison with prior studies of mutual fund flows, we limit the sample to mutual funds with an investment objective of domestic equity.[22] A typical mutual fund name as reported in CRSP is AmSouth Funds: AmSouth Regional Equity Fund; Class A Shares. We begin by dropping

[22] Specifically, we include the following Lipper objective codes: Equity Income (EI), Dedicated Short (S), Hedged (H), Growth and Income (GI), Growth (G), Micro Cap (MR), Small Cap (SG), Mid Cap (MC), and S&P 500 Index (SP). Prior to 1999, we use the corresponding Strategic Insight objective codes.

the share class information to the right of the semicolon. We treat the name left of the colon as the family name and the name right of the colon as the fund name. As with closed-end funds, if a family name is reported, we require that the family name appear in the fund name.

Analogous to our approach with company names, we define fund name fluency as the sum of its *length* score, *Englishness* score, and *dictionary* score. To measure fund name length, we follow the process for company names and drop conjunctions (and, but, etc.) and incorporation terms such as Co., Corp., Inc., and LLC at the end of the name. We also drop ubiquitous words in fund names such as Fund(s) and Portfolio(s). After these adjustments, we count the number of words in the fund name.

Compared with company names, fund names tend to be longer and are more likely to contain nondictionary words. As a result, applying the company fluency score process to fund names without modification would result in relatively little cross-sectional variation in *fluency* score (most funds would be nonfluent). We assume that investors respond to relative name fluency and make adjustments accordingly. Specifically, we assign funds with name lengths below the 25th percentile (four words for closed-end fund and three words for mutual fund) a *length* score of 3, lengths between the 25th and 75th percentile (five words for closed-end fund and four words for mutual fund) a *length* score 2, and funds with name lengths greater than the 75th percentile a *length* score of 1.[23]

For the *dictionary* score, we assign a *dictionary* score of 1 if the proportion of dictionary words in the fund name is greater than the median (67% of words pass the spell-check filter for closed-end funds and 75% for mutual funds) and zero otherwise. Lastly, we continue to use the linguistic algorithm of Travers and Olivier (1978) to assess fund name fluency as defined in subsection 2.2. Specifically, we focus on the lowest *Englishness* score within the fund name and rank funds based on their minimum residual *Englishness* score. Funds in the bottom quintile of Englishness are given an *Englishness* score of 0; all other funds are given an *Englishness* score of 1.

7.2. Control variables and descriptive statistics

For each closed-end fund, we collect a number of additional control variables. We collect the expense ratio and the investment objective from Morningstar. Specifically, Morningstar partitions all closed-end funds into one of the six following investment objectives: Balanced, International Stock, Municipal Bond, Sector Stock, Taxable Bond, and US Stock. We collect monthly net asset values from Compustat and we collect prices, shares outstanding, stock returns, and dividend payouts from CRSP. For each fund-month, we compute closed-end fund premium as log ($Price_{i,t}/NAV_{i,t}$).[24] We refer to negative premiums as discounts. We also compute a number of additional control variables. The full list of variables and the details of their construction are presented in Appendix B.

For each mutual fund, we collect past returns, total net assets, expense ratios and investment objectives from CRSP. We define the monthly net flow into a fund as

$$Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1+R_{i,t})}{TNA_{i,t-1}}, \quad (6)$$

where $R_{i,t}$ is the return on fund i in month t and $TNA_{i,t}$ is the total net asset value for fund i at the end of month t. To minimize the potential impact of errors due to mutual fund mergers or splits (e.g., Elton, Gruber, Blake, 2001), we winsorize flows at the 2.5 and 97.5 percentiles. We again compute a number of additional control variables. The details of their construction are presented in Appendix B.

After all data requirements, the final sample includes 366 closed-end funds and 43,083 closed-end fund—month observations and 7,112 mutual funds and 479,761 mutual fund—month observations. As with company names, the distribution of fund name fluency is bell-shaped. Fourteen percent of closed-end funds have a *fluency* score of 5 and only 3% of funds have a *fluency* score of 1. Similarly, 5% of mutual funds have a *fluency* score of 5 and 2% have a *fluency* score of 1.

The average closed-end fund discount is 4.38%, with a standard deviation of 9.13%. The average monthly mutual fund flow is 1.25%, although the distribution is skewed with median flow being 0.18%. Interestingly, we find that fluent funds tend to be systematically different from nonfluent funds. For example, fluent closed-end funds are bigger and older than nonfluent funds. Specifically, the average fluent closed-end fund (i.e., funds with *fluency* scores greater than 3) has a market equity of $354 million and is 123 months old, and the average nonfluent fund (i.e., funds with *fluency* scores less than 3) has a market equity of $171 million and an age of 93 months.

Similarly, fluent mutual funds are substantially larger and older than non-fluent mutual funds. The average fluent mutual fund has total net assets of 1.3 billion, is 11 months old, and has an average style-adjusted percentile performance rank of 49.09%. In contrast, the average nonfluent fund has $199 million in total net assets, is six months old, and has a performance rank of 50.04%. The systematic differences between fluent and nonfluent funds highlight the importance of examining the relation between fund fluency and investor demand in a regression framework.

7.3. Closed-end fund fluency and discounts

We begin by examining the relation between closed-end fund discounts and fund name fluency by estimating the panel regression

$$Premium_{i,t} = a_0 + a_1 Fluency_{it-1} + \mathbf{a_2}\mathbf{X}_{\text{it}-1} + \varepsilon_{it}, \quad i = 1,\ldots,N \quad t = 1,\ldots,T, \quad (7)$$

where fluency is the fund's aggregate fluency score, $\mathbf{X_{it-1}}$ is a vector of control variables, and ε_{it} is measurement error.[25] Our hypothesis is that funds with fluent names will trade at a

[23] The results are generally not sensitive to the specific choice of length cutoffs.

[24] The results are robust to winsorizing the discount at the 99th and 1st percentile as well as computing the discount as $(Price_{i,t} - NAV_{i,t})/NAV_{i,t}$.

[25] Using Fama and Macbeth instead of panel regression yields very similar results for both closed-end funds and mutual funds.

Table 8
Fund name fluency and closed-end fund premiums.

This table reports the estimates of regressions of closed-end fund premiums over net asset value on fund name fluency and other fund characteristics. The dependent variable is $Log(Price_{i,t}/NAV_{i,t})$ for fund i in month t. *Fluency* score is the sum of the *length*, *Englishness*, and *dictionary* scores. Fund names consisting of less than four, four or five, and greater than five words receive a *length* score of 3, 2, and 1, respectively. Funds in the bottom quintile of Englishness, as measured using a linguistic algorithm, receive an *Englishness* score of 0; all other funds receive an *Englishness* score of 1. Fund names in which more than two-thirds of words satisfy the spell-check filter receive *dictionary* scores of 1; all other funds receive *dictionary* scores of 0. *High* (*Low*) *Fluency* is a dummy variable equal to one if the *fluency* score of the fund is greater than (less than) 3. Definitions for all control variables are presented in Appendix B. The regressions also include investment objective-time fixed effects. Standard errors are clustered by fund, and t-statistics are reported in parentheses.

	1	2	3	4	5
Fluency score	1.04			1.46	0.71
	(3.44)			(3.36)	(2.04)
Length score		0.94			
		(2.16)			
Englishness score		1.76			
		(3.11)			
Dictionary score		0.14			
		(0.21)			
High Fluency			1.43		
			(2.70)		
Low Fluency			−2.04		
			(−2.70)		
Dividend Yield	6.46	6.52	6.49	6.61	5.89
	(8.02)	(8.12)	(8.08)	(6.43)	(5.70)
Expense Ratio	2.33	2.30	2.65	2.47	2.26
	(2.70)	(2.71)	(3.02)	(2.15)	(2.06)
Past Year Return	104.43	103.45	105.19	133.39	57.90
	(6.10)	(6.13)	(6.19)	(6.21)	(2.76)
LN(*Fund Age*)	1.38	1.44	1.44	1.88	0.75
	(2.71)	(2.78)	(2.84)	(2.57)	(1.91)
LN(*Fund Size*)	−0.98	−1.00	−1.01	−1.32	−0.04
	(−3.60)	(−3.70)	(−3.77)	(−2.20)	(−0.10)
Style time dummies	Yes	Yes	Yes	Yes	Yes
R^2	0.253	0.254	0.256	0.319	0.289
Number of observations	47,041	47,041	47,041	23,219	23,822
Sample	All	All	All	Small funds	Large funds

higher premium or smaller discount than less fluent funds and, specifically, that a_1 will be greater than zero. The vector of controls, $\mathbf{X_{it-1}}$, includes several variables to control for differences in agency problems, managerial skill, and arbitrage opportunities. To control for agency problems, we include the fund's expense ratio and dividend payout. Higher expense ratios have a negative impact on the performance of the fund and should result in larger discounts. Higher payout ratios reduce resources under managerial control and could increase the price of the fund relative to NAV.

Funds managed by high ability managers could trade at a smaller discount (e.g., Berk and Stanton, 2007). We include the fund's prior year return as a proxy for managerial ability. Because most funds trade at a discount, funds that are easier to arbitrage are likely to trade at a smaller discount. Following Pontiff (1996), we include fund size as a measure of arbitrage cost. Because funds tend to be issued at a premium, which slowly erodes over time, we also include fund age. In addition, we exclude funds that are less than one year old. Finally, the regressions include investment objective time fixed effects.

Table 8 presents the results of the panel regression, and t-statistics based on standard errors clustered by fund are reported in parentheses. In the first specification, we find that a 1 unit increase in *fluency* score results in a 1.04% increase in the fund premium (or reduction in the discount). This effect is weaker than the 1.90% increase in Tobin's q from Table 4, but it remains economically important and highly significant. In the second specification, we decompose the fluency effect into each of its components. We find that both *length* score and *Englishness* are positively and significantly related to closed-end fund premiums.

In Specification 3, we examine whether fund name fluency has an asymmetric effect on fund discounts. Relative to funds with the middle *fluency* score of 3 (30% of the sample), we find that nonfluent funds with a score of less than 3 trade at a 2.04% greater discount, and fluent funds with a score of greater than 3 trade at 1.43% smaller discount. The slightly stronger effect for nonfluent funds is consistent with the common stock results (see Fig. 1), although the difference between the two estimates is not statistically significant.

We next examine whether the effects are stronger among smaller funds. Each month, we split funds into small and large funds based on the median market value breakpoint. Consistent with the common stock results, we find that the effects of fluency are stronger for smaller funds. Specifically, a 1 unit increase in fluency is associated with a 1.46% increase for small funds versus a 0.71% increase for large funds.

7.4. Mutual fund fluency and fund flows

We next investigate the relation between mutual fund flows and fund name fluency by estimating the panel

Table 9
Fund name fluency and mutual fund flows.

This table reports the estimates of regressions of mutual fund net flows on fund-name fluency and other fund characteristics. The dependent variable is $Flow_{i,t}$ defined as net dollar flow as a percentage of total net assets for fund *i* in month *t*. *Fluency* score is the sum of the *length*, *Englishness* and *dictionary* score. Fund names consisting of less than three, three or four, and greater than four words receive a *length* score of 3, 2, and 1, respectively. Funds in the bottom quintile of Englishness, as measured using a linguistic algorithm, receive an *Englishness* score of 0; all other funds receive an *Englishness* score of 1. Fund names in which more than 75% of words satisfy the spell-check filter receive *dictionary* scores of 1; all other funds receive *dictionary* scores of 0. *High* (*Low*) *Fluency* is a dummy variable equal to one if the *fluency* score of the fund is greater than (less than) 3. Definitions for all other control variables are presented in Appendix B. The regressions also include investment objective time fixed effects. Standard errors are clustered by fund and month, and *t*-statistics are reported in parentheses.

	1	2	3	4	5
Fluency score	0.21			0.23	0.11
	(7.32)			(5.60)	(3.72)
Length score		0.41			
		(9.88)			
Englishness score		0.02			
		(0.49)			
Dictionary score		0.11			
		(1.79)			
High Fluency			0.28		
			(5.73)		
Low Fluency			−0.14		
			(−2.30)		
Low Performance	0.05	0.05	0.05	0.05	0.06
	(12.27)	(12.40)	(12.32)	(8.59)	(11.43)
Mid Performance	0.03	0.03	0.03	0.03	0.03
	(20.17)	(20.19)	(20.18)	(17.91)	(18.17)
High Performance	0.15	0.15	0.15	0.16	0.14
	(21.68)	(21.67)	(21.68)	(18.16)	(18.52)
Volatility	−3.38	−3.09	−3.47	−0.01	−5.92
	(−1.35)	(−1.24)	(−1.39)	(−0.01)	(−2.21)
Log(*Fund Size*)	−0.08	−0.08	−0.08	−0.32	0.06
	(−5.99)	(−6.18)	(−6.03)	(−11.82)	(3.62)
Log(*Fund Age*)	−1.56	−1.58	−1.55	−2.00	−1.24
	(−34.53)	(−35.03)	(−34.42)	(−30.65)	(−26.96)
Expense ratio	−53.67	−54.23	−53.52	−37.17	−67.08
	(−6.18)	(−6.32)	(−6.20)	(−3.36)	(−11.23)
Large Family	0.33	0.37	0.33	0.58	0.01
	(5.74)	(6.37)	(5.73)	(7.88)	(0.17)
Style time dummies	Yes	Yes	Yes	Yes	Yes
R^2	0.164	0.166	0.165	0.165	0.177
Number of observations	479,761	479,761	479,761	239,634	240,127
Sample	All	All	All	Small funds	Large funds

regression

$$Flows_{i,t} = a_0 + a_1 Fluency_{it-1} + \mathbf{a_2}\mathbf{X}_{\mathrm{it-1}} + \varepsilon_{it}, \quad i = 1,...,N, \quad t = 1,...,T, \tag{8}$$

where fluency is the fund's aggregate *fluency* score, $\mathbf{X_{it-1}}$ is a vector of control variables, and ε_{it} is measurement error. Our hypothesis is that funds with fluent names will receive greater fund flows than less fluent funds and, specifically, that a_1 will be greater than zero.

$\mathbf{X_{it-1}}$ includes a number of control variables defined in Appendix B. First, we include controls for past performance. For each fund, we compute its performance over the prior year and create a percentile ranking for the fund relative to all other funds in its investment objective. As prior work shows an asymmetric performance-flow relation, we estimate flows using a piecewise linear regression that allows for different flow-performance sensitivities at different levels of performance (see, e.g., Huang, Wei, and Yan, 2007; and Sirri and Tufano, 1998). Specifically, we create three performance variables: *Low*, *Mid*, and *High*, where *Low*=Min(Rank, 20), *Mid*=Min(0, Rank-*Low*), and *High*=Rank-*Mid*-*Low*. We also include several nonperformance variables that have been shown to influence fund flows. Specifically, we include fund risk as measured by the standard deviation of the fund's return over the past 12 months, fund age (in logs), fund size (in logs), the fund's expense ratio, and a dummy variable for large fund families. Lastly, to control for time-varying demand across certain styles, we include style time fixed effects.

Table 9 presents the results of the panel regression, and *t*-statistics based on standard errors clustered by both fund and time are reported in parentheses. We find that fund name fluency is strongly related to fund flows. A 1 unit increase in fund name fluency increase fund flows by 0.21% per month (or 2.5% per year). This effect is comparable to the increase in flows from moving from the 50th percentile in performance to the 57th percentile. In Specification 2, the decomposition indicates that the *length* score is strongly related to net flows, and there is weaker evidence that the *Englishness* score is also related to flows.

In Specification 3, we find that investors both avoid nonfluent funds and are attracted to fluent funds, although the benefit of a fluent fund name is stronger than the cost of a nonfluent name. This result is at odds with the asymmetric findings for stocks and closed-end funds, but it could be related to greater marketing of non-fluent funds by investment advisers. In particular, we find nonfluent mutual funds tend to have higher 12b–1 fees than fluent funds, which provides greater incentives for them to be recommended by investment advisors.[26]

In Specifications 4 and 5, we split the sample into small funds (funds below the median total net assets for a given date) and large funds. We find that the effect is significant for both small and large funds, but the magnitude is roughly twice as large for smaller funds. Overall, the fund results confirm the impact of name fluency on investor recognition and asset value.

8. Conclusion

Growing evidence exists that investors have a preference for familiar and likeable stocks. For example, investors tend to tilt their portfolio toward locally headquartered stocks and toward companies with large levels of advertising (Coval and Moskowitz, 1999; and Grullon, Kanatas, and Weston, 2004). Investors also gravitate toward Fortune's most admired stocks and shun tobacco stocks (Statman, Fisher, and Anginer, 2008; Hong and Kacperczyk, 2009). In this article, we examine whether the fluency of a company name is another important source of familiarity and affinity that influences investment decisions. Building on the literature in psychology, which finds that fluent stimuli appear more positive and familiar than nonfluent stimuli, and the literature in marketing which emphasizes the importance of product names, we conjecture that investors will have a preference for companies with fluent names.

Consistent with this hypothesis, we find companies with fluent names have higher levels of both retail and mutual fund shareholders as well as greater turnover and smaller transaction price impacts. Moreover, we show that this larger investor base and improved liquidity have important implications for firm value. Specifically, companies with fluent names trade at significant premiums relative to companies with less fluent names.

Our results show a robust positive correlation between name fluency, investor recognition, and firm value. While such a correlation is interesting in its own right, a concern is that name fluency merely proxies for a relevant omitted variable (e.g., management quality). We offer additional evidence that taken together supports the view that name fluency has a causal effect on investor recognition and firm value.

First, when companies undergo names changes they tend to select more fluent names. Although name changes themselves could be endogenous, CEOs often cite fluency-related concepts in motivating the name change. This is consistent with management having an intuitive awareness of the importance of name fluency. Moreover, within-firm variation in fluency (based on name changes) is also positively associated with breadth of ownership, liquidity, and firm value. The results are considerably stronger for the subset of Rebranding name changes in which shifts in the firm's business operations are unlikely. The evidence of within-firm fluency effects helps mitigate the concern than name fluency proxies for an omitted time-invariant firm characteristic.

Second, we find convincing evidence that fluency also impacts other investment decisions. Specifically, we find that fluent mutual funds receive greater fund flows and fluently named closed-end funds trade at smaller discounts from net asset value. Because closed-end fund book values (i.e., net asset value) are based on market prices, they provide a particularly clean test of the impact of fluency on valuation. Taken together, our findings support the view that name fluency has a significant effect on investor recognition and firm value.

Our results suggest a new channel through which companies and investment funds can take advantage of investors' preference for the familiar. Unlike the location of a firm's headquarters, which is likely influenced by economic considerations, or advertising, which can be costly, selecting a fluent name appears to be a relatively low-cost method for improving investor recognition and increasing firm value.

Appendix A

See Table A1.

Appendix B. Description of variables

This appendix describes the construction of the dependent variables and the controls used in the regression analysis. With the exception of *Retail Breadth* and *Retail Turnover*, which span from 1991 to 1996, all other company-level variables are computed each year from 1982 to 2009. Closed-end fund variables and mutual fund variables are computed from 1994 to 2009 and 1992 to 2009, respectively.

Company-level variables

Size—market capitalization computed as share price times total shares outstanding at the end of the year.
Age—number of months since a firm's first return appeared in CRSP.
Book-to-market—book-to-market ratio computed as the book value of equity for the fiscal year ending before the most recent June 30, divided by the market capitalization on December 31 of the same fiscal year.
Volatility—standard deviation of monthly returns during a given year.
Turnover—average monthly turnover (i.e., share volume scaled by shares outstanding) over the 12 months in the year.
Momentum—return on the stock over the past two to 12 months, measured at the end of the year.
NYSE—dummy variable equal to one if the firms trades on the NYSE.
S&P 500—dummy variable equal to one if the firm belongs to the S&P 500.

[26] In untabulated findings, we repeat the mutual fund analysis for the subset of funds that do not charge 12b–1 fees (roughly 10% of the sample; data on 12b–1 fees are often missing) and find nonfluent funds receive −0.61% less in flows and fluent funds receive 0.29% more in flows, which is consistent with the findings for stocks and closed-end funds.

Table A1
Company name *fluency* scores by firm size.

This table reports examples of company names and their *fluency* scores. We report the five smallest and largest companies (based on 2009 market equity) for each fluency group. S1 (L1) reflects the smallest (largest) stock, and S5 (L5) reflects the fifth smallest (largest) stock. For each company, we also report the *length* score, the *dictionary* score, and the *Englishness* score (in that order) in brackets.

	Fluency score				
Group	1 (least fluent)	2	3	4	5 (most fluent)
S1	Helios & Matheson NA [0,0,0]	Manhattan Bridge Capital [0,0,1]	Taitron Components [1,0,1]	Conolog [2,0,1]	Banks.com [2,1,1]
S2	MACC Private Equities [0,0,0]	US Dataworks [1,0,0]	Food Technology Service [0,1,1]	ValueRich [2,0,1]	Multiband [2,1,1]
S3	Nyer Medical Group [0,0,0]	General Employment Enterprises [0,0,1]	Giga-Tronics [1,0,1]	eOn Communications [1,1,1]	Goldfield [2,1,1]
S4	Provident Community Bancshares [0,0,0]	Kent Financial Services [0,0,1]	OccuLogix [2,0,0]	Castle Brands [1,1,1]	Reeds [2,1,1]
S5	TII Network Technologies [0,0,0]	Comstock Homebuilding Companies [0,0,1]	Community Shores Bank [0,1,1]	Zoom Technologies [1,1,1]	Insure.com [2,1,1]
L5	PNC Financial Services Group [0,0,0]	Eli Lilly [1,0,0]	Cisco Systems [1,0,1]	Intel [2,0,1]	Caterpillar [2,1,1]
L4	National Oilwell Varco [0,0,0]	Goldman Sachs Group [0,0,1]	Conocophillips [2,0,0]	Google [2,0,1]	Apache [2,1,1]
L3	EI Du Pont De Nemours [0,0,0]	American International Group [0,0,1]	International Business Machines [0,1,1]	AT&T [2,0,1]	Oracle [2,1,1]
L2	Bristol Myers Squibb [0,0,0]	Johnson & Johnson [1,0,0]	Procter & Gamble [1,0,1]	Microsoft [2,0,1]	Apple [2,1,1]
L1	Freeport McMorRan Copper & Gold [0,0,0]	Wal-Mart Stores [0,0,1]	Exxon Mobil [1,0,1]	General Electric [1,1,1]	Chevron [2,1,1]

Illiquidity—Amihud (2002) measure computed using all daily data available for a given calendar year.

Advertising/Sales (*R&D/Sales*)—total advertising expenditures (research and development expenditures) scaled by total sales. Following Himmelberg, Hubbard, and Palia (1999) we set missing values of *Advertising/Sales* and *R&D/Sales* to zero and include an indicator variable that equals one when there is a missing value and zero otherwise.

Strong Brand—dummy variable equal to one if the firm is ever ranked among the top one hundred global brands by Interbrand (2001–2010), or the top five hundred global brands by Brandirectory (2007–2010). We use forward-looking information on brand ranking as a conservative control.

Profitability—EBITDA scaled by book value of assets. We set negative values of profitability to zero and include an indicator variable that equals one when there is a negative value and zero otherwise.

Growth—sales growth measured over the past three years. If less than three years of sales data are available, sales growth is estimated using all available data. If no information on prior sales is available, we set sales growth to zero and include an indicator variable that equals one when there is a missing value and zero otherwise.

Leverage—book value of debt scaled by book value of assets.

Asset Turnover—sales divided by book value of assets.

Payout—sum of dividends and repurchases divided by net income.

Tobin's q—enterprise value (debt plus market value of equity) scaled by book value (debt plus book value of equity).

Ownership—a measure of the number of investors holding the firm stocks. We consider two measures of *Ownership. Mutual Fund Breadth*—number of unique mutual funds holding the firms' stock at the end of the given year. The number of mutual fund shareholders is computed from the Thomson Financial s12 files. *Retail Breadth*—number of retail investors holding the firm's stock at the end of the given year. The number of retail shareholders is taken from a large discount brokerage that contains the holdings of 78 thousand households from January 1991 to November 1996.

Retail Turnover—average monthly retail turnover over the 12 months in the year and is also computed using the discount brokerage data set.

Closed-end fund variables

Size—market capitalization computed as share price times total shares outstanding in the prior month.

Age—number of months since a fund's first return appeared in CRSP.

Past Return—average monthly return on the fund over the prior year (measured each year).

Expense Ratio—annual expense ratio as reported in Morningstar, winsorized at the 99th percentile.
Dividend Yield—total dividends paid by the fund over the past year scaled by the funds' net asset value at the end of the year.

Mutual fund variables

Size—total net assets in the prior month.
Age—number of months since a fund's first return appeared in CRSP.
Rank—percentile ranking of a fund's return over the prior 12 months within its respective investment objective category: *Low*=Min (Rank, 20), *Mid*=Min (60, Rank-*Low*), *High*=*Rank-Low-Mid*.
Volatility—standard deviation of a fund's return over the past 12 months.
Expense Ratio—annual expense ratio as reported in CRSP.
Big Family—dummy variable equal to one if the fund belongs to a family in top quintile of total number of funds offered.

References

Alter, A., Oppenheimer, D., 2006. Predicting short-term stock fluctuations by using processing fluency. Proceedings of the National Academy of Sciences 103, 9369–9372.
Alter, A., Oppenheimer, D., 2009. Uniting the tribes of fluency to form a metacognitive nation. Personal Social Psychological Review 13, 219–235.
Amihud, Y., 2002. Illiquidity and stock returns: cross section and time series effects. Journal of Financial Markets 5, 31–56.
Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17, 223–249.
Bao, Y., Shao, A., Rivers, D., 2008. Creating new brand names: effects of relevance, connotation, and pronunciation. Journal of Advertising Research 48, 148–162.
Barber, B., Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. Journal of Finance 55, 773–806.
Barber, B., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Quarterly Journal of Economics 116, 261–292.
Battalio, R., Mendenhall, R., 2005. Earnings expectations, investor trade size, and anomalous returns around earnings announcements. Journal of Financial Economics 77, 289–319.
Berk, J., Stanton, R., 2007. Managerial ability, compensation, and the closed-end fund discount. Journal of Finance 62, 529–556.
Benartzi, S., 2001. Excessive extrapolation and the allocation of 401(k) accounts to company stock. Journal of Finance 56, 1747–1764.
Bodnaruk, A., Ostberg, P., 2009. Does investor recognition predict returns? Journal of Financial Economics 91, 208–226.
Chen, H., Noronha, G., Singal, V., 2004. The price response to S&P 500 index additions and deletions: evidence of asymmetry and a new explanation. Journal of Finance 59, 1901–1929.
Cohen, L., 2009. Loyalty-based portfolio choice. Review of Financial Studies 22, 1213–1245.
Cooper, I., Kaplanis, E., 1994. Home bias in equity portfolios, inflation hedging and international capital market equilibrium. Review of Financial Studies 7, 45–60.
Cooper, M., Dimitrov, O., Rau, R., 2001. A rose.com by any other name. Journal of Finance 56, 2371–2388.
Cooper, M., Gulen, H., Rau, R., 2005. Changing names with style: mutual fund name changes and their effects on fund flows. Journal of Finance 60, 2825–2858.
Coval, J., Moskowitz, T., 1999. Home bias at home: local equity preference in domestic portfolios. Journal of Finance 54, 2045–2073.
Coval, J., Moskowitz, T., 2001. The geography of investment: informed trading and asset prices. Journal of Political Economy 109, 811–841.
Edmans, A., Garcio, D., Norli, O., 2007. Sports sentiment and stock returns. Journal of Finance 62, 389–406.
Edmans, A., Goldstein, I., Jiang, W., 2012. The real effects of financial markets: the impact of prices on takeovers. Journal of Finance 67, 933–971.
Elton, E., Gruber, M., Blake, C., 2001. A first look at the accuracy of CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund database. Journal of Finance 56, 2415–2430.
Falkenstein, E., 1996. Preferences for stock characteristics as revealed by mutual fund portfolio holdings. Journal of Finance 51, 111–135.
Fama, E., French, K., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.
Fama, E., French, K., 2008. Dissecting anomalies. Journal of Finance 63, 1653–1678.
Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.
Fong, K., Holden, C., Trzcinka, C., 2011. What are the best liquidity proxies for global research? Unpublished working paper. University of New South Wales, Kensington, Australia.
French, K., Poterba, J., 1991. Investor diversification and international equity markets. American Economic Review 81, 222–226.
Frieder, L., Subrahmanyam, A., 2005. Brand perceptions and the market for common stock. Journal of Financial and Quantitative Analysis 40, 57–85.
Gompers, P., Metrick, A., 2001. Institutional investors and equity prices. Quarterly Journal of Economics 116, 229–259.
Green, T.C., Hwang, B.H., 2012. IPOs as lotteries: skewness preference and first-day returns. Management Science 58, 432–444.
Grinblatt, M., Keloharju, M., 2001. How distance, language, and culture influence stockholdings and trades. Journal of Finance 56, 1053–1073.
Grullon, G., Kanatas, G., Weston, J., 2004. Advertising, breadth of ownership, and liquidity. Review of Financial Studies 17, 439–461.
Himmelberg, C., Hubbard, G., Palia, D., 1999. Understanding the determinants of managerial ownership and the link between ownership and performance. Journal of Financial Economics 53, 353–384.
Hirsh, B., Seaks, T., 1993. Functional form in regression models of Tobin's *q*. Review of Economics and Statistic 75, 381–385.
Hirshleifer, D., Shumway, T., 2003. Good day sunshine: stock returns and the weather. Journal of Finance 58, 1009–1032.
Hoberg, G., Phillips, G., 2010a. Product market synergies and competition in mergers and acquisitions: a text-based analysis. Review of Financial Studies 23, 3773–3811.
Hoberg, G., Phillips, G., 2010b. Text-based network industries and endogenous product differentiation. Unpublished working paper. University of Maryland, College Park, MD.
Hong, H., Kacperczyk, M., 2009. The price of sin: the effect of social norms on stock markets. Journal of Financial Economics 93, 15–36.
Huang, J., Wei, K., Yan, H., 2007. Participation costs and the sensitivity of fund flows to past performance. Journal of Finance 62, 1273–1311.
Huberman, G., 2001. Familiarity breeds investment. Review of Financial Studies 14, 659–680.
Ivkovic, Z., Weisbenner, S., 2005. Local does as local is: information content of the geography of individual investors' common stock investments. Journal of Finance 60, 267–306.
Kadlec, G., McConnell, J., 1994. The effect of market segmentation and liquidity on asset prices: evidence from exchange listings. Journal of Finance 49, 611–636.
Kaul, A., Mehrotra, V., Morck, R., 2000. Demand curves for stocks do slope down: new evidence from an index weights adjustment. Journal of Finance 55, 893–912.
Kumar, A., 2009. Who gambles in the stock market? Journal of Finance 64, 1889–1933.
Labroo, A., Dhar, R., Schwarz, N., 2008. Of frowning watches and frog wines: semantic priming, perceptual fluency, and brand evaluation. Journal of Consumer Research 34, 819–831.
Lang, L., Stulz, R., 1994. Tobin's *q*, diversification, and firm performance. Journal of Political Economy 102, 1248–1280.
Lee, C., Shleifer, A., Thaler, R., 1991. Investor sentiment and the closed-end fund puzzle. Journal of Finance 46, 75–109.
Lesmond, D., Ogden, J., Trzcinka, C., 1999. A new estimate of transactions costs. Review of Financial Studies 12, 1113–1141.
Massa, M., Simonov, A., 2006. Hedging, familiarity, and portfolio choices. Review of Financial Studies 19, 633–685.
McGlone, M., Tofighbakhsh, J., 2000. Birds of a feather flock cojointly? Rhyme as reason in aphorisms. Psychological Science 11, 424–428.
Merton, R.C., 1987. Presidential address: a simple model of capital market equilibrium with incomplete information. Journal of Finance 42, 483–510.

Newey, W., West, K., 1987. A simple positive definite heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Oppenheimer, D., 2006. Consequences of erudite vernacular utilized irrespective of necessity: problems with using long words needlessly. Applied Cognitive Psychology 20, 139–156.
Pontiff, J., 1996. Costly arbitrage: evidence from closed-end funds. Quarterly Journal of Economics 111, 1135–1151.
Rubin, D., Friendly, M., 1986. Predicting which words get recalled: measures of free recall, availability, goodness, emotionality, and pronunciability for 925 nouns. Memory and Cognition 14, 79–94.
Sarkissian, S., Schill, M., 2004. The overseas listing decision: new evidence of proximity preference. Review of Financial Studies 17, 769–809.
Schwarz, N., Bless, H., Strack, F., Klump, G., Rittenauer-Schatka, H., Simmons, A., 1991. Ease of retrieval as information: another look at the availability heuristic. Journal of Personality and Social Psychology 61, 195–202.
Seasholes, M., Zhu, N., 2010. Individual investors and local bias. Journal of Finance 65, 1987–2010.
Shah, A., Oppenheimer, D., 2007. Easy does it: the role of fluency in cue weighting. Judgment and Decision Making 2, 371–379.
Shleifer, A., 1986. Do demand curves for stocks slope down? Journal of Finance 41, 579–590.
Sirri, E., Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance 53, 1589–1622.
Song, H., Schwarz, N., 2009. If it's difficult to pronounce, it must be risky: fluency, familiarity, and risk perception. Psychological Science 20, 135–138.
Statman, M., Fisher, K., Anginer, D., 2008. Affect in a behavioral asset pricing model. Financial Analyst Journal 64, 20–29.
Travers, J., Olivier, D., 1978. Pronounceability and statistical 'Englishness' as determinants of letter identification. American Journal of Psychology 91, 523–538.
Tversky, A., Kahneman, D., 1973. Availability: a heuristic for judging frequency and probability. Cognitive Psychology 5, 207–232.
Wu, Y., 2010. What's in a name? What leads a firm to change its name and what the name foreshadows. Journal of Banking and Finance 34, 1344–1359.