The power of primacy:

Alphabetic bias, investor recognition, and market outcomes

Heiko Jacobs and Alexander Hillert*

This draft: November 2013

Abstract

Research has revealed that the convention of alphabetical name ordering tends to provide an advantage to those positioned in the beginning of the alphabet. This paper is the first to explore implications of this form of alphabetic bias in the natural setting of financial markets. We reveal empirically that the bias leaves many discernible traces in equilibrium outcomes. For instance and all else equal, stocks with names early in alphabet have 5% to 15% higher trading activity and lower costs of trading. These phenomena are strongest for firms disproportionately traded by individual investors.

Keywords: Trading behavior, behavioral finance, name effects, limited attention, ordering effects

JEL Classification Codes: G02, G12, G14

---

*Heiko Jacobs, Finance Department, University of Mannheim, L 5, 2, 68131 Mannheim. E-Mail: jacobs@bank.bwl.uni-mannheim.de. Phone: +49 621 181 3453. Fax:+49 621 181 1534. Alexander Hillert, Chair of International Finance and CDSB, University of Mannheim, L9, 1-2, 68131 Mannheim. E-Mail: hillert@bwl.uni-mannheim.de. Phone: +49 621 181 1462. Fax:+49 621 181 1664. Parts of this research project have been conducted while the first author visited the School of Banking & Finance at the University of New South Wales. We would like to thank Russell Jame, Markku Kaustia, Alok Kumar as well as seminar participants at the Third Helsinki Finance Summit on Investor Behavior, at the 20th Annual Meeting of the German Finance Association (DGF), and at the University of Mannheim for helpful comments.

1 Introduction

Being on or near the top of a list is in many ways important in academia. For example, Brogaard et al. (2012) uncover that articles placed first, second, and third in an issue of top finance and economic journals enjoy all else equal about 50%, 26%, and 17% more future citations than articles published near the back of an issue. The authors conclude that it is not completely clear whether this placement effect reflects editors' private information about article quality or readers' limited attention. Other effects, however, point more directly to the latter explanation. Often these phenomena are strongly related to the widespread convention of alphabetization, for example with regard to credits on coauthored publications in economics. Einav and Yariv (2006) and van Praag and van Praag (2008) demonstrate that, due to the increased visibility of first-named authorship, researchers with initials closer to the beginning of the alphabet benefit by various measures of academic success. These even seem to include better chances of receiving tenure at top ten economics departments or prestigious awards such as the Nobel Prize.

There are more examples for this form of "alphabetic bias" in academia. Richardson (2010) reveals that in the case of a well-established medical imaging journal, reviewers whose last name started with an A received almost twice as many review invitations as their colleagues towards the end of the alphabet. As reportedly several thousands of other publications, the journal relied on software which automatically presented a typically alphabetically sorted list of potential reviewers. Editors seemed to "almost certainly" (p. 215) exhibit the natural tendency to invite the first panel of appropriate reviewers. A closely related mechanism applies to positions as visiting academics (Merritt (1999)). Moreover, Meer and Rosen (2011) document a comparatively low probability of alumni with names located at the end of the alphabet donating to their university. This appears to be driven by volunteers, who personally call potential donors in alphabetical order, and often simply run out of time before they reach the end.

As the convention of alphabetization is a fact of everyday life, "alphabetic bias" is clearly far from being limited to academia.[1] Collectively, this leads "The Economist" to arrive at the (hyperbolical) conclusion: "Over the past century, all kinds of unfairness and discrimination have been denounced or made illegal. But one insidious form continues to thrive: alphabetism. This [...] refers to discrimination against those whose surnames begin with a letter in the lower half of the alphabet." (Economist (2001), p.13).

This research project is to the best of our knowledge the first to study the role of alphabetization in the specific context of financial markets. We thereby particularly concentrate on the stock market, which offers a natural and promising setting for several reasons.

First, investment decisions in stock markets are economically important, both with respect to investor performance and the efficient allocation of capital across firms. Second, the setting offers a rich set of data on market participants' behavior, control variables, and aggregated outcomes, enabling us to distinguish more precisely among alternative hypotheses than some existing empirical studies in other settings do.

Third, due to the thousands of stocks available, investors face a considerable search problem (e.g. Barber and Odean (2008), Merton (1987)), making ordering effects likely to occur. Fourth, there is a widespread convention of listing company names alphabetically in newspapers, indices, data bases, and most other sources of information. Figure 1 provides motivating examples from popular websites.

Please insert figure 1

Bloomberg.com allows to browse through stocks by the first letter of the firm name (on top of the page) or by sectors. Any subsample selected by one of these methods yields an

[1]For instance, it drives the chance of getting access to over-subscribed public services (Jurajda and Münich (2010)) or the likelihood of receiving votes on alphabetically ordered ballots (e.g. Bakker and Lijphart (1980), Miller and Krosnick (1998), Ho and Imai (2008)).

alphabetically sorted list. Russel.com presents the official list of Russell 3000 members on 37 pages with alphabetical ordering of firm names. Similarly, finance.yahoo.com presents members of the Nasdaq Composite on 49 pages. For any market segment, nytimes.com offers firm variables like size or nearness to the 52 week high, but again firms with names early in the alphabet are displayed on the first pages.

Against this background, we test the hypothesis that firms towards the beginning of such lists enjoy privileged treatment. Put simply, we comprehensively explore stock-market implications of the intuitive conjecture of "The Economist" (2001, p. 13): "It has long been known that a taxi firm called AAAA cars has a big advantage over Zodiac cars when customers thumb through their phone directories."[2] Our findings, whose essence is captured in figure 2, support this claim.

Please insert figure 2

The graph compares trading activity (as measured by monthly turnover) and costs of trading (as measured by the monthly Amihud (2002) illiquidity ratio) for NYSE/AMEX firms grouped by sorting on their firm name in alphabetical order. The benchmark in figure 2 are firms whose name is above the 75th percentile in a given month. Findings reveal that firms positioned earlier in the alphabet appear to be strikingly more recognized by investors. For instance, firms in the top 5% have about 13% higher turnover and 11% lower costs of trading than firms in the lowest 25% of the alphabet. Stock characteristics as well as actual trading data from a large discount brokerage suggest that these phenomena tend to be strongest for firms with a high fraction of individual investors.

Importantly, the findings illustrated in figure 2 are estimates from multivariate regres-

[2] This claim is backed up by eye movement data (e.g. Lohse (1997)). Moreover, as Einav and Yariv (2006) highlight (p. 186): "In fact, this sort of influence on attention appears to be heavily exploited in the realm of advertising. For example, the 2003-2004 Los Angeles Westside Yellow Pages reveal more than 450 listed business with names containing a seemingly redundant initial A, as in "A-Approved Chimney Services," "A Any Way Bail Bonds," "A Budget Moves," and the like."

sions over a time period ranging from January 1995 to December 2009, which control for 35 firm characteristics. These include, for instance, market capitalization, industry membership, analyst coverage, earnings surprises, national and local media coverage, different types of firm news extracted from 8-K filings, or company name fluency (Green and Jame (2013)).

We find similar results across firm universes, sample periods, model specifications, and econometric approaches. Results also carry over to the breadth of ownership, as measured by the total number of shareholders. We also find some (albeit weaker) evidence for the impact of a firm's alphabetical rank on firm valuation as well as on coverage by analysts and the local media. Moreover, the analysis of name changes is also consistent with the idea of "alphabetic bias" affecting investor decision making.

Finally, we turn to the mutual fund market, which offers another promising setting. Again, investors face multiple investment alternatives with many sources of information presenting lists of funds in alphabetical order. Thus, funds with names early in the alphabet may be comparatively more likely to be part of the choice set of investors facing time and capacity constraints. Moreover, the mutual fund business is a trillion dollar market, making investment decisions economically important. Finally, as the literature review below shows, recent work has uncovered the impact of name effects on mutual fund flows, rendering the ordering of names another important issue. Our empirical analysis of monthly fund flows from 1992 to 2012 supports this conjecture. The effect of alphabetic fund name ordering is concentrated in small funds for which search costs tend to be higher. In this subgroup, funds within the top of the alphabet enjoy about 2.5% to 3.5% more inflows p.a., holding all other factors fixed.

As browsing from the top of lists is natural human behavior, it seems likely that the underlying mechanisms of our findings are similar to the ones proposed for the academic settings sketched above. Comparable to the case of papers with multiple authors, being named early on a stock list might boost visibility. Comparable to the case of review invitations and visiting positions, top-ranked stocks might also have a high chance of already satisfying the needs of market participants in search of investment alternatives. As in the case of alumni donations and the other settings, investors might also run out of time, capital, interest, or face other constraints before they reach the bottom of a stock list.

At least for some investors, several other more subtle effects may be at work. For instance, being among the first items on a list may (potentially unconsciously) often be associated with superior quality, such as in the example of finance papers placed first, second, and third in an issue. Confirming evidence is brought forward by Ang et al. (2010) who uncover that investors exhibit affect for class A shares relative to class B shares.

Moreover, and similar to the impact of first-named authorship, a stock's permanent visibility may eventually translate into a higher degree of familiarity among investors, which in turn has been shown to be an important determinant of financial decision making (e.g. Huberman (2001), Grinblatt and Keloharju (2001)). Finally, there is a large amount of literature on the primacy effect, which Carney and Banaji (2012) characterize in their literature review as follows (p.1): "What is experienced first is remembered better (...), it drives attachment more strongly (...), creates stronger association with the self (...), influences impressions more decisively (...), and persuades more effectively (...)." Thus, even if market participants worked through a number of potential investments from the top to the bottom of e.g. a typical stock index, they may well eventually choose one of the stocks listed first.

Our study contributes to several strands of literature. First, it adds to the emerging research on name effects in financial markets. This work shows that factors unrelated to fundamentals can influence investor behavior and potentially also market outcomes. In a clinical study, Rashes (2001) uncovers excess comovement between two stocks with very similar ticker symbols and attributes this to confusion among predominantly small investors. Vivid examples for name effects are presented in Cooper et al. (2001) and Cooper et al. (2005b): During (after) the internet bubble, firms that simply changed their name to (away from) “dot.com names” enjoyed positive abnormal post-announcement returns.

Similarly, cosmetic name changes among mutual funds yield positive abnormal inflows without delivering better performance (Cooper et al. (2005a)). Green and Jame (2013) show that funds with more fluent names attract higher inflows. Two recent contribution uncover name effects also at the fund manager-level (Niessen-Ruenzi and Ruenzi (2013), Kumar et al. (2013)): stereotypes associated with female or foreign-sounding fund manager names lead, all else equal, to substantially lower inflows. Our contribution is to show that even the ordering of names, and not the meaning of the names themselves, influences equilibrium outcomes, both in the stock market and the mutual fund market.

Second, our findings add to the understanding of the drivers of stock-level trading activity. While trading activity is often closely linked to patterns in liquidity or return predictability (see e.g. Hong and Stein (2007)), the sources of its large cross-sectional differences are still not well known (e.g. Chordia et al. (2007), Lo and Wang (2000), or Statman et al. (2006)). Several papers have progressed on this front by uncovering that specific variables deemed to be related to investor recognition and familiarity are positively correlated with stock-level turnover. Loughran and Schultz (2004) and Jacobs and Weber (2012) explore the role of firm location. Grullon et al. (2004) focus on advertising expenditures, and Chen et al. (2004) concentrate on index addition effects.

The study most closely related to ours in this respect is Green and Jame (2013). They argue that firms with short, easy to pronounce, rememberable names (such as Google or Apple) attract, all else equal, more investors, which translates into higher stock-level turnover, lower transaction price impact, and higher firm value. We later control for the drivers of trading volume studied in all of these papers. Our contribution is to show that a previously neglected and seemingly minor detail, essentially the first letter of a company name, has a surprisingly strong and pervasive impact on trading activity, liquidity, and breadth of ownership.

Third, our study adds to the vibrant literature on limited investor attention. Due to the complexity of financial markets on the one hand and cognitive constraints on the other hand, market participants have to be selective with regard to information processing. As a consequence, investors cannot take all alternatives into account simultaneously (e.g. Kahneman (1973), Hirshleifer et al. (2009)). Both theory and evidence suggests that investors instead may resort to category thinking or other forms of complexity reducing heuristics, and thereby tend to neglect potentially value-relevant information.[3] Our contribution is to present a novel channel through which (primarily retail) investor attention can be assessed.

In essence, we uncover a previously neglected factor which might influence investor behavior. As the convention on alphabetical ordering is widespread, these findings might have implications for several actors in financial markets. For instance, they might be of relevance to exchange-listed firms in search of comparatively cheap methods to increase liquidity and breadth of ownership. With regard to the presentation and marketing of their product universe, results might be of interest to e.g. index providers and mutual fund companies. Findings might also be relevant for investors searching for neglected stocks.

[3] A non-exhaustive list of reference papers (in alphabetical order...) includes Barber and Odean (2008), Barberis and Shleifer (2003), Barberis et al. (2005), Boyer (2011), Cohen and Frazzini (2008), DellaVigna and Pollet (2009), Greenwood (2008), Green and Hwang (2009), Menzly and Ozbas (2010), Mullainathan (2002), Peng (2005), Peng and Xiong (2006), and Peress (2008).

2 Empirical approach and baseline findings

Our goal is to test for the existence of an "alphabetic bias" in the context of investment decisions, which might be strong enough to affect economic aggregates. Thus, referring back to the example sketched above, explanatory variables which quantify the difference between "AAAA cars" and "Zodiac cars" are of primary interest.[4]

We build on two measures, which are both based on the time-series of (historical and current) company names as provided by CRSP. The first measure simply indicates the relative position of a firm's name in the firm universe under consideration (e.g. NYSE/AMEX firms or Nasdaq firms). In each month, we sort all eligible firms on their official company name in alphabetically ascending order. The variable, in the following referred to as *position continuous* and used in the baseline analysis, is then computed as the rank of the firm divided by the number of all firms. Thus, *position continuous* is uniformly distributed over the interval]0,1].

Our second measure takes potential non-linearities into account. The impact of alphabetization might be disproportionately strong for firms at the beginning of a list, and then decay gradually. To capture this line of reasoning mathematically, we introduce several disjunct dummies, in the following referred to as *position dummies*. *Position dummy* 5 takes on a value of one if the company name is among the first 5% in a given month and zero otherwise. *Position dummy* 25 is 1 if the firm name is between percentiles 5 and 25. Finally, *position dummy* 50 (*position dummy* 75) is 1 if the name is between percentiles 25 and 50 (50 and 75). Consequently, findings are benchmarked against the firms near the end of the alphabet (i.e., percentile 75 and above).

[4]As in this example, we only consider firms with names starting with a letter. This is intended to assure a clean, conservative analysis. About 0.15% of monthly observations in our baseline sample represent firms for which the name starts with a number according to CRSP. Examples include 1st Source Corporation or 3Com. However, findings are very robust with respect to the inclusion of these firms.

To comprehensively control for other variables, the empirical analysis combines data from several sources. The CRSP/Compustat Merged Database provides common stock as well as balance sheet information. Among others, we augment this data set with analyst data obtained from I/B/E/S, extensive local and national firm-level media coverage gathered from LexisNexis, and all 8-K filings compiled from EDGAR. The employed variables are described in the following. Details on their construction are provided in the appendix.

The dependent variables in our baseline analysis are stock-level measures of trading activity and trading costs. With respect to the first, we focus on firm turnover as it "yields the sharpest empirical implications and is the most natural measure" (Lo and Wang (2000), p. 12). We logarithmize turnover as it is naturally skewed. As our measure of the costs of trading, we rely on the logarithmized Amihud (2002) illiquidity ratio. The ratio is computed as the monthly average of daily illiquidity measures, defined as the absolute daily stock return divided by daily trading volume in millions of dollars.

Due to the differences in market structure and the well-documented double-counting issue for stocks trading on Nasdaq (e.g. Atkins and Dyl (1997), Anderson and Dyl (2005)), one needs to be careful when comparing trading across markets. We therefore run our analysis separately for NYSE/AMEX and Nasdaq stocks. Similar patterns across these disjunct subsamples would moreover suggest that our findings represent a more generalized phenomenon rather than an effect of exchange-specific factors or elaborate data mining. To additionally estimate an overall effect across exchanges, we follow the method suggested in e.g. Loughran and Schultz (2005) by doubling trading volume for NYSE and AMEX stocks, pooling them with Nasdaq stocks, and adding a Nasdaq dummy to the regression.

We start with the framework proposed in Chordia et al. (2007), which aims at comprehensively identifying cross-sectional determinants of expected trading activity. Their econometric approach involves running predictive Fama and MacBeth (1973)-type regressions of a measure of trading activity for stock i in month t+1 ($Trading_{i,t+1}$) on up to n sets of lagged explanatory firm characteristics ($\sum_{k=1}^{n} Set_{k,i,t}$).

$$Trading_{i,t+1} = \gamma_{0,t} + \sum_{k=1}^{n} \gamma_{k,t} Set_{k,i,t} + \epsilon_{i,t+1} \quad (1)$$

$\gamma_{k,t}$ refers to the vector of coefficients obtained for the control variables of set k in month t. We adopt and extend the two sets of control variables considered in Chordia et al. (2007). In total, we consider two dependent variables (logarithmized turnover, logarithmized illiquidity ratio), three samples (NYSE/AMEX, Nasdaq, all) and three different models (with different controls and sample starts ranging from 1963 to 1995). To be in line with the approach of Chordia et al. (2007), statistical inference in these 18 specifications is based on heteroskedasticity- and autocorrelation-consistent (HAC) Newey and West (1987) t-statistics. We later rely on yearly data as well as on predictive panel or firm fixed effect regressions to test the robustness of this baseline approach.

The first set of explanatory variables in Chordia et al. (2007) has the advantage of long data availability so that the beginning of the sample period can be set to July 1963 for AMEX and NYSE stocks. In comparison, the sample period starts in January 1983 for Nasdaq stocks, when trading volume data for these stocks becomes comprehensively available in CRSP. We refer to this setting as model 1.

More specifically, we include two return-related variables, which might proxy for rebalancing needs (e.g. Chordia et al. (2007)) or investor recognition (e.g. Merton (1987), Barber and Odean (2008)). The two variables are defined as the stock return in the previous month if positive (negative) and zero otherwise. This distinction is motivated by possible asymmetric effects caused by short-selling constraints or the disposition effect (e.g. Odean (1998), Grinblatt and Keloharju (2001)). We employ the book-to-market ratio to account for differences in fundamentals and in investor attention between value and growth stocks. We include lagged logarithmized versions of the firm's market capitalization, age, and nominal share price to control for potential differences in ownership structure, visibility, or trading costs (Falkenstein (1996), Kumar and Lee (2006), Merton (1987)).[5] Following Chordia et al. (2007) and the references therein, we also use a firm's leverage ratio and beta.

For each model specification, we only consider stock months for which all control variables as specified above are available. Untabulated findings show that this leads to our sample firms being slightly larger than the standard CRSP common stock universe. As we reveal later, ordering effects tend to be stronger for small firms so that our approach might be considered conservative. Moreover, in all three stock samples (NYSE/Amex, Nasdaq, all), the correlation of the five position measures (*position continuous*, four *position dummies*) with any of the other explanatory variables in model 1 is between -0.03 and 0.03. Thus, with respect to standard firm characteristics, firms with names at the beginning of the alphabet do not seem to be systematically different from firms with names at the end of the alphabet.

Panel A of table 1 (2) presents main findings with regard to name ordering-induced cross-sectional differences in turnover (illiquidity).

Please insert tables 1 and 2

Several findings are noteworthy. First, the number of firms which meet data requirements for model 1 is large. In the specification with all three exchanges, about 3,300 firms enter the analysis in a typical month. Thus, ordering effects might play a role in the decision of which stock to trade.

Second, *position continuous* indeed matters for the cross-section of both trading activity and the costs of trading. For NYSE/AMEX (Nasdaq) stocks since 1963 (1983), the implied turnover increase for firms at the beginning of the alphabet relative to their

[5] These three variables tend to be substantially correlated, but our findings are very robust to specifications in which we only include one or two of them.

counterparts at the end is all else equal about 11% (5.5%). Across exchanges, it is more than 8.5%. For the Amihud (2002) illiquidity ratio, findings are similar. Findings suggest that firms listed early in the alphabet have about 7.5% to 9% lower trading costs. These values are not only highly statistically significant, but also economically meaningful.

Third, the impact of alphabetic ranking on both turnover and illiquidity is quite robust across time and can be identified in about 72% to 90% of all sample months. We now test whether the effect remains persistent when including more comprehensive control variables.

For model specification 2, we start by considering the same additional variables that Chordia et al. (2007) use in their full specification. We use the lagged logarithmized number of analysts following a stock as a proxy for information diffusion, the mass of informed agents and the firm's visibility (e.g. Brennan and Subrahmanyam (1995), Hong et al. (2000)). We use analyst forecast dispersion as a proxy for differences of opinion (e.g. Diether et al. (2002)). Finally, estimation uncertainty about fundamental values is quantified by lagged measures of earnings surprises and earnings volatility.

While this specification is equivalent to model 2 in Chordia et al. (2007), we augment the model (for month t+1) with additional variables shown to be related to trading activity in further studies. We start with return-related variables: we employ the return twelve months ago (t-11) as well as the cumulative return over t-7 to t-1 to account for effects related to (seasonal) momentum as discussed in e.g. Heston and Sadka (2008), Statman et al. (2006) or Lee and Swaminathan (2000). Moreover and inspired by e.g. Seasholes and Wu (2007) or Huddart et al. (2009), a 52-week high dummy is considered.

We continue by including variables intended to better control for nonlinearities in size as well as for index effects (e.g. Chen et al. (2004)). To this end, we consider the squared value of the lagged (logarithmized) market capitalization, and dummies for S&P

500 membership, Dow Jones 30 membership (for NYSE/AMEX stocks), and Nasdaq 100 membership (for Nasdaq stocks). We include idiosyncratic volatility, computed as the standard deviation from the residual of rolling regressions of a firm's excess return on a Carhart (1997) model. The alpha from this regression is included as it has been argued to contain a premium related to liquidity or heterogeneous information (e.g. Lo and Wang (2000)). Grullon et al. (2004) show that advertising expenditures are positively related to investor recognition. Findings in e.g. Green and Jame (2013) suggest that research and development expenditures (R&D) might have a similar effect. We use both variables and scale them by total sales.

Next, we control for geographic factors related to trading activity. Loughran and Schultz (2005) show that firms located in urban areas are traded heavier due to the larger local investor base. Following their approach, we construct dummies for firms in urban and rural regions (see also Hong et al. (2008)). We also follow their analysis in adding the logarithmized number of employees to the regression. Lastly, we use a set of dummies to control for the 49 Fama and French (1997) industries.

Model 2 relies partly on lagged monthly analyst (and other) data, which is not reliably available before 1980 (e.g. Hong et al. (2000)). Therefore, Model 2 starts in February 1980 for the NYSE/AMEX sample and again in January 1983 for Nasdaq stocks.

Main findings for turnover (illiquidity) are displayed in panel B of table 1 (2). The impact of *position continuous* remains very stable, thus inferences do not change. Again, the impact of alphabetic name ranking on turnover (illiquidity) is estimated to range roughly from 6% to 9% (7% to 8%).

The major goal of model 3 is to control more comprehensively for firm-specific news and investor recognition, both of which are likely to be positively related to stock-level trading activity and liquidity. A number of studies suggest that accounting for media coverage is a powerful approach in this context (e.g. Barber and Odean (2008), Engelberg and Parsons (2011), Peress (2008), Tetlock (2011)). Recent research (e.g. Gurun and Butler (2012)) has uncovered that local media are of particular importance. We thus gather in total roughly 261,000 firm-specific articles in four leading national newspapers (New York Times, USA Today, Wall Street Journal, and Washington Post) as well as about 888,000 articles in 41 local newspapers. We construct two control variables as the logarithmized number of previous month articles in national and local newspapers, respectively. We also add state dummies to the regression. The latter are intended to account for cross-sectional differences in local media availability as well as in local economic conditions, which have been shown to affect local firms' liquidity (Bernile et al. (2013)).

In addition, we gather in total about 421,000 8-K filings which are individually accessible via the EDGAR database. The SEC[6] defines an 8-K filing as a "'current report' companies must file with the SEC to announce major events that shareholders should know about". For instance, potential events include changes in executive management or ownership, entry into a material definitive agreement, matters related to accountants and financial statements, or asset redeployments. Thus, 8-K filings offer a very rich data set to control for firm news. Along pre-defined dimensions, firms have to specify the type of reportable event(s) in the main body of an 8-K filing. We thus write a parsing algorithm to identify the exact type of news. On this basis, we construct in total ten variables with the logarithmized number of different firm events in the previous month.

8-K filings are available from 1995 onwards, and coverage for many local newspapers also does not start before the middle of the 90ties. The sample period for model 3 thus starts in February 1995.

Finally, we use data provided by Green and Jame (2013) to also control for the recently discovered impact of company name fluency. However, untabulated tests show that

[6] http://www.sec.gov/answers/form8k.htm

name fluency and alphabetic rank clearly represent distinct phenomena. The correlation between *position continuous* and the fluency score ranges from roughly -0.01 to about 0.03, depending on the time period and firm universe. Lagged company name fluency is available until December 2009, which consequently marks the end of the sample period for model 3.

Main findings for turnover (illiquidity) in this full model specification are displayed in panel C of table 1 (2). It turns out that the effect is even slightly more pronounced than in models 1 and 2. For instance, in the combined NYSE/AMEX/Nasdaq specification, the implied increase in turnover (decrease in trading costs) for firms positioned at the beginning of the alphabet is about 9% (9.5%), holding the 35 firm characteristics in model 3 constant.

In sum, the analysis so far reveals a strong and pervasive impact of alphabetically sorted firm names on both trading activity and liquidity. The effect can be identified across firm universes, sample periods, and models. We now turn to further sensitivity checks and additional analyses.

3 Robustness checks and further insights

Alternative frequencies and alternative econometric approaches As insights from the following five tests are similar across models and in order to conserve space, table 3 only reports the results for the full regression specification (model 3).

Please insert table 3

We start by changing the econometric approach to mitigate the impact of potentially unobserved firm heterogeneity not captured by our comprehensive control variables. As a first approach, we run panel regressions with year dummies and standard errors double-clustered by firm and month. Petersen (2009) argues that double-clustering produces correctly sized standard errors, regardless of whether a potential firm effect is permanent or temporary. As a second approach, we include both firm-fixed effects and month-fixed effects. Thus, the coefficient for *position continuous* is based on within-firm variation over time. We later explicitly study name changes.

As names tend to be a rather stable firm characteristic, one might argue that relying on yearly (as opposed to monthly) data might be a useful approach. We thus replicate the baseline Fama and MacBeth (1973)-type analysis as well as the two alternative econometric methods with low-frequency yearly data. To this end, we reconstruct all independent and dependent variables by relying on their yearly average (e.g. for turnover) or end of year values (e.g. for industry classification).

As table 3 uncovers, neither the use of conservative econometric approaches nor the change of the sample frequency materially changes the main insights. In almost any setting and for both turnover (panel A) and liquidity (panel B), coefficients remain highly statistically significant and economically meaningful.

Turnover for Nasdaq stocks poses an exception. With firm-fixed and month-fixed (or year-fixed) effects, the impact of *position continuous* is estimated to be only around 3%, which is not significant anymore. For both NYSE/Amex stocks and the combined firm universe, however, the effect remains large (around 10%) and highly significant.

With regard to illiquidity, findings for all firm samples (including Nasdaq firms) tend to be even slightly larger than in the baseline analysis. All else equal, estimates suggest a liquidity difference of 8% to 16% between firms with names early in the alphabet and firms with names late in the alphabet.

In the overall picture, inferences from the baseline analysis thus remain unchanged.

Impact of outliers One might be concerned that our findings may be driven by outliers. For each month separately, we therefore winsorize turnover and the illiquidity ratio at the 99% level. Rerunning all regressions verifies that (untabulated) findings remain virtually unchanged.

Alternative measure of ordering Table 4 reports findings when we rely on disjunct *position dummies*, as described in section 2, to measure the relative position of a firm's name. We run predictive panel regressions and report standard errors which are double-clustered by firm and month. In line with the idea of investors browsing from the top of lists of alphabetically sorted firm names, firms at the very beginning of the alphabet seem to be disproportionately more recognized by investors than all other firms.

For instance, estimates from the full model specification across exchanges indicate that the first 5% of firms have, all else equal, 10.7% higher trading activity and 11.7% lower costs of trading than their counterparts with names in percentiles 75 and above. The effect becomes smaller, apparently in a non-linear fashion, for firms later in the alphabet. For instance, while firms with names in percentiles 5 to 25 still enjoy 8.2% higher turnover and 6.2% lower costs of trading, there is hardly any effect for firms with names in percentiles 50 to 75. In sum, findings strongly support the insights from the baseline analysis.

Please insert table 4

Alternative measure of firm identification We have replicated the analysis with ticker symbols instead of firm names. Findings turn out to be very similar. This is not surprising as in roughly 95% of all firm months, the first letter of the company name equals the first letter of the ticker symbol. Therefore, relative positions in the alphabet tend to be very similar for the vast majority of observations.

Alternative firm universe In their stock selection process, market participants might not consider the whole stock universe, but apply preceding filter rules. Such screens may be related to stock exchanges, as in our baseline analysis. Arguably another popular way of restricting the firm universe is to focus on specific industries. To explore the idea that alphabetization-induced ordering effects should also be identifiable in these subgroups, we rerun the analysis for seven broad sectors as defined by SIC divisions.[7] Dependent variables are again either turnover or illiquidity. Moreover, we again consider models 1 to 3 as well as both turnover and illiquidity. The appendix reports the main findings from pooled panel regressions with double-clustered standard errors.

In 95% of the in total 42 regressions, the coefficient of *position continuous* goes in the predicted direction: trading activity is larger and costs of trading are lower for firms whose names show up early on an alphabetically sorted list of stocks belonging to a given industry. Findings are also economically meaningful: about 90% of the regressions estimate the impact of *position continuous* to be at least 5%. Moreover, results are fairly robust across model specifications. In sum, findings strongly support the insights from the baseline analysis.

Alternative approach: name changes As they allow us to focus on within-firm variation, company name changes offer a conceptually different setting to examine the role of alphabetic name ordering. In an attempt to balance the trade-off between a sufficiently large sample size and a comprehensive set of control variables, we rely on model 2 for the NYSE/Amex/Nasdaq universe from the baseline analysis. The model covers the sample period from January 1983 to December 2011. We only consider firm name changes which go along with a change in *position continuous* of at least 0.01, i.e. imply an increase or decrease in the relative ranking of at least 1%.

[7] We only consider those industry divisions which consist of at least 40 eligible stocks in each month of the respective sample period. This cut-off is somewhat arbitrarily set and meant to restrict the analysis to larger industry groups for which ordering effects are more likely to play a role.

The direction of these changes does not seem to be completely random. In total, we are able to identify 1,259 name changes. Out of these, 57.11% (42.89%) lead to a position earlier (later) in the alphabetical ranking, which is a highly significant imbalance. We then run pooled panel regressions with all control variables of model 2, including month-fixed effects. However, we also introduce firm-fixed effects as well as a dummy which quantifies the period after name changes. Moreover, we use contemporaneous (instead of lagged) controls to timely capture effects related to the name change.

Our central result is that the insights from the between-firm baseline analysis also hold for the within-firm analysis in the context of name changes. With turnover (the Amihud (2002) measure) as dependent variable, the coefficient of *position continuous* is -0.132 (0.129) and significant at the one percent level. These findings again address concerns that our baseline analysis might be driven by an omitted variable problem.

Other market participants: analysts and the media It is worth noting that our regressions so far included controls for analyst coverage, national and local media coverage. We may now ask ourselves whether analysts and the media themselves might exhibit a tendency to cover firms which are, due to their early position in the alphabet, attention-grabbing. To explore this hypothesis, we run panel regressions with double-clustered standard errors, thereby concentrating on the full model (specification 3 and Amihud illiqudity as additional control) and the whole stock universe.

The data partly supports the hypothesis. In the case of the number of analysts as dependent variable, *position continuous* takes on a value of -0.025 (t-statistic -1.48), indicating that firms named early in the alphabet receive ceteris paribus about 2.5% more coverage than firms at the bottom of the alphabet. There is no effect for the national media. However, in the case of local media, the coefficient is -0.031 with a t-statistic of -1.94. The effect is strongest for the first 5% of firms in the alphabet, which enjoy about 5% more local media coverage.

Breadth of ownership and firm valuation Firms at the beginning of the alphabet are traded more often and at lower costs. If these findings are indicative of higher investor recognition, then we might expect these firms to have also a higher breadth of ownership, holding all else equal. To explore this relationship, we run panel regressions similar to the ones in table 3, but we now use the logarithmized total number of shareholders as our dependent variable. As this value is only available on a yearly basis, we run yearly regressions for all models. We also include the Amihud illiqudity ratio as additional control. Panel A of table 5 shows the findings for the NYSE/AMEX/Nasdaq universe.

Please insert table 5

Findings are in line with our expectations: regardless of whether we use *position continuous* or *position dummies*, firms listed early in the alphabet appear to have ceteris paribus about 7% to 9% more shareholders.

Finally, previous work (e.g. Merton (1987)) suggests a positive relation between breadth of ownership and firm valuation. We use the logarithmized market-to-book ratio as a proxy for a firm's valuation and run regressions as specified above. Panel B of table 5 provides some supportive results. While the effect seems weaker than in most of the other tests, firms with names at the very beginning of the alphabet still seem to trade at about 5% higher valuations.

Investor types To the extent that our baseline findings are rooted in cognitive constraints of market participants, they might be most pronounced in stocks with a high fraction of individual investors. This subgroup is widely considered to exhibit more biases and to have less resources and knowledge than professional investors (see e.g. Barber and Odean (2013) or Grinblatt and Keloharju (2001)).

We explore the impact of cross-sectional differences in the investor base in two ways. First, we focus on subgroups of stocks known to be a natural habitat of less sophisticated investors. Second, we directly study individual investors' trading decisions and portfolio holdings.

The trading of individual investors has been shown to be concentrated in small stocks (e.g. Kumar and Lee (2006), Dorn et al. (2008)), in stocks with lottery-type features (Kumar (2009)), in stocks covered by the media (e.g. Barber and Odean (2008), Engelberg et al. (2012)), and in stocks with low institutional ownership (e.g. Lee et al. (1991), Lemmon and Portniaguina (2006)). We thus partition our sample along these dimensions.

We rely on the most comprehensive specification (i.e., NYSE/AMEX/Nasdaq universe, model 3) although inferences do not depend on this choice. We divide the sample into small and large stocks based on the median market capitalization in the previous month. As a second approach, we distinguish between lottery and non-lottery stocks, for which we follow the approach developed by Kumar (2009). Lottery firms (non-lottery firms) are firms with above (below) median idiosyncratic volatility, above (below) median idiosyncratic skewness, and below (above) median nominal share price. A third approach relies on (national) press coverage. As media coverage is highly positively correlated with firm size (e.g. Fang and Peress (2009)), we first orthogonalize the monthly number of articles with respect to lagged market capitalization before we perform a median split. Finally, the fourth approach distinguishes firms with above and below median lagged institutional ownership. Data is gathered from the Thomson-Reuters Institutional Holdings (13F) Database. For these four subsamples, we then run panel regressions as in table 3. Panel A of table 6 shows the main findings.

Please insert table 6

As expected, the effect of alphabetization is considerably stronger for stocks which are likely to be disproportionately traded by less sophisticated investors. Most notably, small firms (firms with low institutional ownership) positioned early in the alphabet are estimated to have all else equal about 14% (17%) higher turnover than corresponding firms late in the alphabet, while the corresponding estimate for large firms (firms with high institutional ownership) is less than 7% (3%). The findings for lottery stocks and national media also go in the predicted direction, but are less in magnitude.

In the overall picture, findings are even stronger in the case of illiquidity. Firms which have a name within the top of the alphabet and are at the same time likely to be a natural investment habitat of individual investors are about 14% to 19% more liquid than corresponding firms at the end of the alphabet. For large firms, non-lottery firms, firms neglected by the media, or firm with a high fraction of institutional investors, the effect is about 1.5% to 9% and in most cases statistically not distinguishable from zero.

We next turn to study the behavior of individual investors directly, and contrast their trading patterns with the overall market. To this end, we compute retail breadth of ownership and retail turnover by using data from a large discount broker over 1991 to 1996 (see Barber and Odean (2000), Barber and Odean (2001), and Kumar (2009) for more information). Due to the early sample period, only model 1 or 2 come into question for the analysis. Panel B of table 6 shows the findings for both specifications.

The coefficients on *position continuous* suggest that firms early in the alphabet enjoy roughly 13% to 16% more retail trading activity. In contrast, the corresponding estimates for unconditional trading activity over the same time period are about 8% to 9%. Similarly, findings indicate that firms early in alphabet have about 5% to close to 8% higher breadth of retail ownership, whereas the corresponding estimates for unconditional breadth of ownership are 4% to close to 6% during that time. In sum, individual investors appear to suffer more from "alphabetic bias" than other market participants.

Experimental evidence on ordering effects Our interpretation of the findings so far rests on the assumption that investors facing a number of investment alternatives concentrate more on the alternatives offered first. Examples for potential channels behind such ordering effects are given in the introduction: search satisfaction, various constraints (e.g. with regard to time, attention, capital), primacy effect, and other factors.

To additionally validate our conjecture on ordering effects, the appendix reports insights from an experiment designed to mirror the presentation format in popular stock information sources (see e.g. the examples in figure 1). The experimental setting also allows us to isolate the impact of pure ordering effects, holding firm and investor characteristics fixed.

More specifically, we present subjects a four-page list of unnamed stocks in randomized order, together with actual firm characteristics known to influence trading behavior. We ask subjects to indicate which stocks they would like to trade. Each stock in the first half of the list has a virtually identical twin in the second half, which however is not known to subjects. In the absence of ordering effects, the likelihood of being selected should be very similar for both constituents of a given pair. However, the twin presented first receives far more attention than the twin towards the end of the list. In general, findings reveal strong ordering effects in stock selection. Combining this phenomenon with the quasi-norm of alphabetization is likely to result in an "alphabetic bias" in investment decisions. In sum, insights from the experiment are thus closely in line with the results from the empirical analysis of the stock market.

4 Another setting: mutual fund flows

The mutual fund industry offers another promising setting in which to study the determinants of investor behavior. As in the stock market, investors have to choose from a large set of alternatives. This holds true even if one narrows down the decision to well-defined market segments, such as open-end U.S. equity mutual funds with a domestic investment focus, as we do in the analysis below. Averaged across all months in our sample from January 1992 to December 2012, there are close to 1,500 funds which satisfy all data requirements. Thus, search costs are likely to play a role in the investment process. Indeed, Sirri and Tuffano (1998) test and verify the hypothesis that "consumers would purchase those funds that are easier or less costly for them to identify" (p. 1607). A substantial body of empirical work further provides support for this idea of cognitive overload in the mutual fund investment decision.[8]

Moreover, the mutual fund market is very large, making capital allocations economically important. In the average month of our sample, the combined assets under management total $1.69 trillion. Thus, if "alphabetic bias" does play a role in this large and liquid market, then this type of ordering effect is likely to matter in general for financial decision making.

For the empirical analysis, we rely on the CRSP survivorship bias free mutual fund database. The isolation of ordering effects requires a relatively homogeneous and at the same time large cross-section. We therefore restrict our focus to open-end U.S. domestic equity mutual funds. We use the third and fourth character of the CRSP style code to assign each fund month to one of in total 19 styles in our sample.[9] We use parsing algorithms (e.g. Gil-Bazo and Ruiz-Verdu (2009)) and manual screens to exclude index funds. We moreover drop very small funds that never have net total assets of over $5 million during the sample period. Finally, we exclude observations if any of the control

[8]A non-exhaustive list includes Barber et al. (2005), Cooper et al. (2005a), Haslem (2012), Hortacsu and Syverson (2004), Jain and Wu (2000), Koehler and Mercer (2009), and Solomon et al. (2013).

[9]The CRSP style codes combine the information contained in Wiesenberger Objective Codes (1962-1993), Strategic Insight Objective codes (1993-1998), and Lipper Objective codes (from 1998 on), all of which have been widely used in previous work. The most common style classifications are "small-cap funds", "mid-cap funds", "growth funds", "growth and income funds", "equity income funds", "technology sector funds" and "real-estate sector funds".

variables (see below) cannot be computed. To avoid multiple counting, we aggregate all share classes of the same fund using MFLinks (e.g. Daniel et al. (1997), Kacperczyk et al. (2008)). The sample period begins in January 1992, when all (partly lagged) variables relied on are available in the desired frequency.

The dependent variable is the net inflow or outflow for fund i in month t. Following the consensus in the literature, a fund flow is defined as the percentage change in total net assets ($TNA_{i,t-1}$) which is not driven by the fund's return net of fees ($ret_{i,t}$):[10]

$$Fund\ flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}}{TNA_{i,t-1}} - ret_{i,t} \quad (2)$$

The independent variables of interest are again measures which indicate the position of a given fund name in the universe of eligible funds. Consequently, we rely on *position continuous* as well as on *position dummies* defined as for the analysis in the stock market. We rely on the fund name reported by CRSP.[11]

The selection of control variables is guided by the large amount of previous work on the determinants of fund flows. The computation of these variables is explained in detail in table 7. At the fund-level, we consider the following characteristics: size, age, expense ratio, turnover ratio, past return and risk, and number of share classes. To control for the well-established convex relationship between fund performance and fund flows (e.g. Chevalier and Ellison (1997), Barber et al. (2005)), we take additional style-specific performance measures into account. Specifically, for each year and each of the 19 investment objectives separately, we determine the relative performance position of each fund. The resulting

[10]In order to mitigate the impact of outliers and to be closer to the normality assumption, we winsorize the distribution at the 5% and 95% level in each month separately.

[11]Typical names are "AAL Capital Growth Fund", "Buffalo Small Cap Fund", or "Navellier Performance Funds: Navellier Aggressive Micro Cap Portfolio". Sometimes, CRSP reports a "the" in front of the actual fund name. A typical example is "The Glenmede Fund, Inc.: Glenmede Large Cap Value Portfolio". In these cases, we drop the "the", which however has little impact on our findings.

performance rank variable is evenly distributed between 0 and 1. We allow for a non-linear effect by also including the squared value of the performance rank. Finally, as a parsimonious approach inspired by Green and Jame (2013), we compute the length of the fund name defined as the number of letters, after dropping share-class information and incorporation terms when applicable.

At the fund family-level, we compute the total assets under management as well as the number of offered funds. We also include a dummy for the three largest families in each month. At the style-level, we compute growth rates from the value-weighted flows of all funds with the same investment objective in a given month. In addition, all regressions include style-year fixed effects. Standard errors are double-clustered by fund and month.

In untabulated findings, we have verified that the position measures are weakly correlated with any of the other explanatory variables. Moreover, there is virtually no relation between position measures and future fund returns. Thus, the regression setting seems well-suited to isolate the impact of pure ordering effects induced by alphabetization of fund names. Table 7 shows the main findings.

Please insert table 7

We run regressions separately for all funds (models 1 and 2) as well as for small funds (models 3 and 4) and large funds (models 5 and 6), determined by a monthly median split. Model 1 reveals that the impact of *position continuous* is as expected: funds with names at the beginning of the alphabet generate ceteris paribus about 0.13% higher inflows each month. This effect is statistically significant and comparable to moving from the 50th percentile of average mutual fund flows to the 53rd percentile.

Models 3 to 6 uncover that the effect is driven by small funds with names at the very beginning of the alphabet. This is in line with the limited attention hypothesis as small funds are likely to be less visible than large funds (see e.g. the argumentation in Sirri and Tuffano (1998)). Thus, these funds might benefit disproportionately from being listed near the top of a list. In general, the impact of position measures is only significant in the subsample of smaller than median funds. Among those, the first 5% of funds benefit by far the most. As model 4 shows, they achieve about 0.315% higher inflows per month (or about 3.8% higher inflows per year) than funds towards the end of the alphabet. This highly significant effect is equivalent of moving from the 50th percentile of the distribution of flows for below median-sized sample funds to the 57th percentile. Even in this subsample, the average fund still has about $80 million under management. Thus, findings are important from an economic perspective. The appendix shows that inferences remain qualitatively unchanged if we use alternative ways to control for style, fund family, and time effects.

In the overall picture, our insights from the stock market can thus be transferred to the mutual fund market. Being listed early in the alphabet appears to increase investor recognition, which eventually leaves discernable traces in economic aggregates.

5 Conclusion

Sorting names alphabetically is an omnipresent convention. For many settings such as academia, economics, or politics, this widespread practice has been shown to yield an advantage to those positioned early in the alphabet. This paper is the first to analyze this type of "alphabetic bias" in financial markets.

We find that a higher alphabetic ranking provides stocks with higher share turnover, investors with lower transaction costs, and firms among others with broader ownership.

Digging deeper, we find these effects to be most pronounced for stocks disproportionately traded by individual or otherwise less sophisticated investors. The cross-section of mutual fund flows lends further support to the idea that alphabetical ordering matters for investor decision making.

These novel findings might have implications for several actors in financial markets. For instance, in an attempt to affect consumer choice, many firms are willing to pay for their products to be displayed in an attention-grabbing matter (e.g. Amstrong et al. (2009)). In the context of the stock market as well as the mutual fund market, our findings reveal that a seemingly minor detail such as essentially the first letter of a name can have a similar impact free of charge.

References

Amihud, Y., 2002, "Illiquidity and stock returns: Cross-section and time-series effects," *Journal of Financial Markets*, 5, 31–56.

Amstrong, M., J. Vickers, and J. Zhou, 2009, "Promince and consumer search," *RAND Journal of Economics*, 40, 209–233.

Anderson, A.-M., and E. A. Dyl, 2005, "Market structure and reported trading volume: NASDAQ versus the NYSE," *Journal of Financial Research*, 28, 115–131.

Ang, J., A. Chua, and D. Jiang, 2010, "Is A better than B? How affect influences the marketing and pricing of financial securities," *Financial Analysts Journal*, 66, 40–54.

Atkins, A. B., and E. A. Dyl, 1997, "Market structure and reported trading volume: NASDAQ versus the NYSE," *Journal of Financial Research*, 20, 291–304.

Bakker, E., and A. Lijphart, 1980, "A crucial test of alphabetic voting: the elections at the University of Leiden, 19731978," *British Journal of Political Science*, 10, 521–525.

Barber, B., and T. Odean, 2001, "Boys will be boys: gender, overconfidence, and common stock investment," *Quarterly Journal of Economics*, 116, 261–292.

Barber, B. M., and T. Odean, 2000, "Trading is hazardous to your wealth," *Journal of Finance*, 55, 773–806.

——— , 2008, "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *Review of Financial Studies*, 21, 785–818.

——— , 2013, "The behavior of individual investors," *Handbook of the Economics of Finance*, Elsevier, edited by Constantnides, Harris, and Stulz, 1533–1569.

Barber, B. M., T. Odean, and L. Zheng, 2005, "Out of sight, out of mind: The effects of expenses on mutual fund flows," *Journal of Business*, 78, 2095–2120.

Barberis, N., and A. Shleifer, 2003, "Style investing," *Journal of Financial Economics*, 68, 161–199.

Barberis, N., A. Shleifer, and J. Wurgler, 2005, "Comovement," *Journal of Financial Economics*, 75, 283–317.

Bernile, G., G. Korniotis, A. Kumar, and Q. Wang, 2013, "Local business cycles and local liquidity," *Unpublished working paper, University of Miami, University of Michigan.*

Boyer, B. H., 2011, "Style related comovement: Fundamentals or labels?," *Journal of Finance*, 66, 307–332.

Brennan, M. G., and A. Subrahmanyam, 1995, "Investment analysis and price formation in securities markets," *Journal of Financial Economics*, 38, 361–381.

Brogaard, J., J. Engelberg, and C. A. Parsons, 2012, "Networks and productivity: Causal evidence from editor rotations," *Journal of Financial Economics, forthcoming.*

Carhart, M. M., 1997, "On persistence in mutual fund performance," *Journal of Finance*, 52, 57–82.

Carney, D. R., and M. R. Banaji, 2012, "First is best," *PLoS ONE*, 7, 1–5.

Chen, H., G. Noronha, and V. Singal, 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.

Chevalier, J., and G. Ellison, 1997, "Risk taking by mutual funds as a response to incentives," *Journal of Political Economy*, 105, 1167–1200.

Chordia, T., S.-W. Huh, and A. Subrahmanyam, 2007, "The cross-section of expected trading activity," *Review of Financial Studies*, 20, 709–740.

Cohen, L., and A. Frazzini, 2008, "Economic links and predictable returns," *Journal of Finance*, 63, 1977–2011.

Cooper, M. J., O. Dimitrov, and R. Rau, 2001, "A rose.com by any name," *Journal of Finance*, 56, 2371–2388.

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

Cooper, M. J., A. Khorana, I. Osobova, A. Patel, and P. R. Rau, 2005b, "Managerial actions in response to a market downturn: valuation effects of name changes in the dot.com decline," *Journal of Corporate Finance*, 11, 319–335.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, "Measuring mutual fund performance with characteristic-based benchmarks," *Journal of Finance*, 52, 1035–1058.

DellaVigna, S., and J. M. Pollet, 2009, "Investor inattention and Friday earnings announcements," *Journal of Finance*, 64, 709–749.

Diether, K. B., C. J. Mallow, and A. Scherbina, 2002, "Differences of opinion and the cross section of stock returns," *Journal of Finance*, 57, 2113–2141.

Dorn, D., G. Huberman, and P. Sengmueller, 2008, "Correlated trading and returns," *Journal of Finance*, 63, 885–920.

Economist, T., 2001, "As easy as ZYX," *Vol. 360, No. 8237, p. 13.*

Einav, L., and L. Yariv, 2006, “What’s in a surname? The effects of surname initials on academic success,” *Journal of Economic Perspectives*, 20, 175–187.

Engelberg, J., and C. A. Parsons, 2011, “The causal impact of media in financial markets,” *Journal of Finance*, 66, 67–97.

Engelberg, J., C. Sasseville, and J. Williams, 2012, “Market madness? The case of mad money,” *Management Science*, 58, 351–364.

Falkenstein, E. G., 1996, “Preferences for stock characteristics as revealed by mutual fund holdings,” *Journal of Finance*, 51, 111–135.

Fama, E. F., and K. R. French, 1997, “Industry costs of equity,” *Journal of Financial Economics*, 43, 153–193.

Fama, E. F., and J. D. MacBeth, 1973, “Risk, return, and equilibrium: Empirical tests,” *Journal of Political Economy*, 81, 607–636.

Fang, L., and J. Peress, 2009, “Media coverage and the cross-section of stock returns,” *Journal of Finance*, 64, 2023–2052.

Gil-Bazo, J., and P. Ruiz-Verdu, 2009, “The relation between price and performance in the mutual fund industry,” *Journal of Finance*, 64, 2153–2183.

Green, T. C., and B.-H. Hwang, 2009, “Price-based return comovement,” *Journal of Financial Economics*, 93, 37–50.

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

Greenwood, R., 2008, “Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights,” *Review of Financial Studies*, 21, 1153–1186.

Grinblatt, M., and M. Keloharju, 2001, “How distance, language, and culture influence stockholdings and trades,” *Journal of Finance*, 56, 1053–1073.

Grullon, G., G. Kanatas, and J. P. Weston, 2004, “Advertising, breadth of ownership, and liquidity,” *Review of Financial Studies*, 17, 439–461.

Gurun, U. G., and A. W. Butler, 2012, “Don’t believe the hype: Local media slant, local advertising, and firm value,” *Journal of Finance*, 67, 561–597.

Haslem, J. A., 2012, “Mutual funds and individual investors: Advertising and behavioral issues,” *Unpublished working paper, University of Maryland.*

Heston, S. L., and R. Sadka, 2008, "Seasonality in the cross-section of stock returns," *Journal of Financial Economics*, 87, 418–445.

Hirshleifer, D., S. S. Lim, and S. H. Teoh, 2009, "Driven to distraction: Extraneous events and underreaction to earnings news," *Journal of Finance*, 64, 2289–2325.

Ho, D. E., and K. Imai, 2008, "Estimating causal effects of ballot order from a randomized natural experiment," *Public Opinion Quarterly*, 72, 216–240.

Hong, H. G., J. D. Kubik, and J. C. Stein, 2008, "The only game in town: Stock-price consequences of local bias," *Journal of Financial Economics*, 90, 20–37.

Hong, H. G., T. Lim, and J. C. Stein, 2000, "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies," *Journal of Finance*, 55, 265–295.

Hong, H. G., and J. C. Stein, 2007, "Disagreement and the stock market," *Journal of Political Economy*, 36, 3–27.

Hortacsu, A., and C. Syverson, 2004, "Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds," *Quarterly Journal of Economics*, 119, 403–456.

Huberman, G., 2001, "Familiarity breeds investment," *Review of Financial Studies*, 14, 659–680.

Huddart, S., M. Lang, and M. H. Yetman, 2009, "Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence," *Management Science*, 55, 16–31.

Jacobs, H., and M. Weber, 2012, "The trading volume impact of local bias: Evidence from a natual experiment," *Review of Finance*, 16, 867–901.

Jain, P. C., and J. S. Wu, 2000, "Truth in mutual fund advertising: Evidence on future performance and fund flows," *Journal of Finance*, 55, 937–958.

Jurajda, S., and D. Münich, 2010, "Admission to selective schools, alphabetically," *Economics of Education Review*, 29, 1100–1109.

Kacperczyk, M., C. Sialm, and L. Zheng, 2008, "Unobserved actions of mutual funds," *Review of Financial Studies*, 21, 2379–2416.

Kahneman, D., 1973, "Attention and effort," *Prentice Hall, New Jersey.*

Koehler, J. J., and M. Mercer, 2009, "Selection neglect in mutual fund advertisements," *Management Science*, 55, 1007–1121.

Kumar, A., 2009, "Who gambles in the stock market?," *Journal of Finance*, 64, 1889–1933.

Kumar, A., and C. C. Lee, 2006, "Retail investor sentiment and return comovements," *Journal of Finance*, 61, 2451–2486.

Kumar, A., A. Niessen-Ruenzi, and O. Spalt, 2013, "What is in a name? Mutual fund flows when managers have foreign sounding names," *Unpublished working paper, University of Miami, University of Mannheim, Tilburg University.*

Lee, C. M., A. Shleifer, and R. H. Thaler, 1991, "Investor sentiment and the closed-end fund puzzle," *Journal of Finance*, 46, 75–109.

Lee, C. M., and B. Swaminathan, 2000, "Price momentum and trading volume," *Journal of Finance*, 55, 20172069.

Lemmon, M., and E. Portniaguina, 2006, "Consumer confidence and asset prices: some empirical evidence," *Review of Financial Studies*, 19, 1499–1529.

Lo, A. W., and J. Wang, 2000, "Trading volume: Definitions, data analysis, and implications of portfolio theory," *Review of Financial Studies*, 13, 257–300.

Lohse, G. L., 1997, "Consumer eye movement pattern on yellow pages advertising," *Journal of Advertising*, 26, 61–73.

Loughran, T., and P. Schultz, 2004, "Weather, stock returns, and the impact of localized trading behavior," *Journal of Financial and Quantitative Analysis*, 39, 343–364.

——— , 2005, "Liquidity: Urban versus rural firms," *Journal of Financial Economics*, 78, 341–374.

Meer, J., and H. S. Rosen, 2011, "The ABCs of charitable solicitation," *Journal of Public Economics*, 95, 363371.

Menzly, L., and O. Ozbas, 2010, "Market segmentation and cross-predictability of returns," *Journal of Finance*, 65, 1555–1580.

Merritt, D. J., 1999, "Calling professor AAA: how to visit at the school of your choice," *Journal of Legal Education*, 49, 557–563.

Merton, R. C., 1987, "A simple model of capital market equilibrium with incomplete information," *Journal of Finance*, 42, 483–510.

Miller, J. M., and J. A. Krosnick, 1998, "The impact of candidate name order on election outcomes," *Public Opinion Quarterly*, 62, 291–330.

Mullainathan, S., 2002, "Thinking through categories," *Unpublished working paper, MIT.*

Newey, W. K., and K. D. West, 1987, "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

Niessen-Ruenzi, A., and S. Ruenzi, 2013, "Sex matters: Gender and prejudice in the mutual fund industry," *Unpublished working paper, University of Mannheim.*

Odean, T., 1998, "Are investors reluctant to realize their losses?," *Journal of Finance*, 53, 1775–1798.

Peng, L., 2005, "Learning with information capacity constraints," *Journal of Financial and Quantitative Analysis*, 40, 307–329.

Peng, L., and W. Xiong, 2006, "Investor attention, overconfidence and category learning," *Journal of Financial Economics*, 80, 563–602.

Peress, J., 2008, "Media coverage and investors' attention to earnings announcements," *Unpublished working paper, Insead.*

Petersen, M. A., 2009, "Estimating standard errors in finance panel data sets: Comparing approaches," *Review of Financial Studies*, 22, 435–480.

Rashes, M. S., 2001, "Massively confused investors making conspicuously ignorant choices (MCI-MCIC)," *Journal of Finance*, 56, 1911–1927.

Richardson, M. L., 2010, "Alphabetic bias in the selection of reviewers for the American Journal of Roentgenology," *American Journal of Roentgenology*, 191, 213–216.

Seasholes, M. S., and G. Wu, 2007, "Predictable behavior, profits, and attention," *Journal of Empirical Finance*, 14, 590–610.

Sirri, E. R., and P. Tuffano, 1998, "Costly search and mutual fund flows," *Journal of Finance*, 53, 1589–1622.

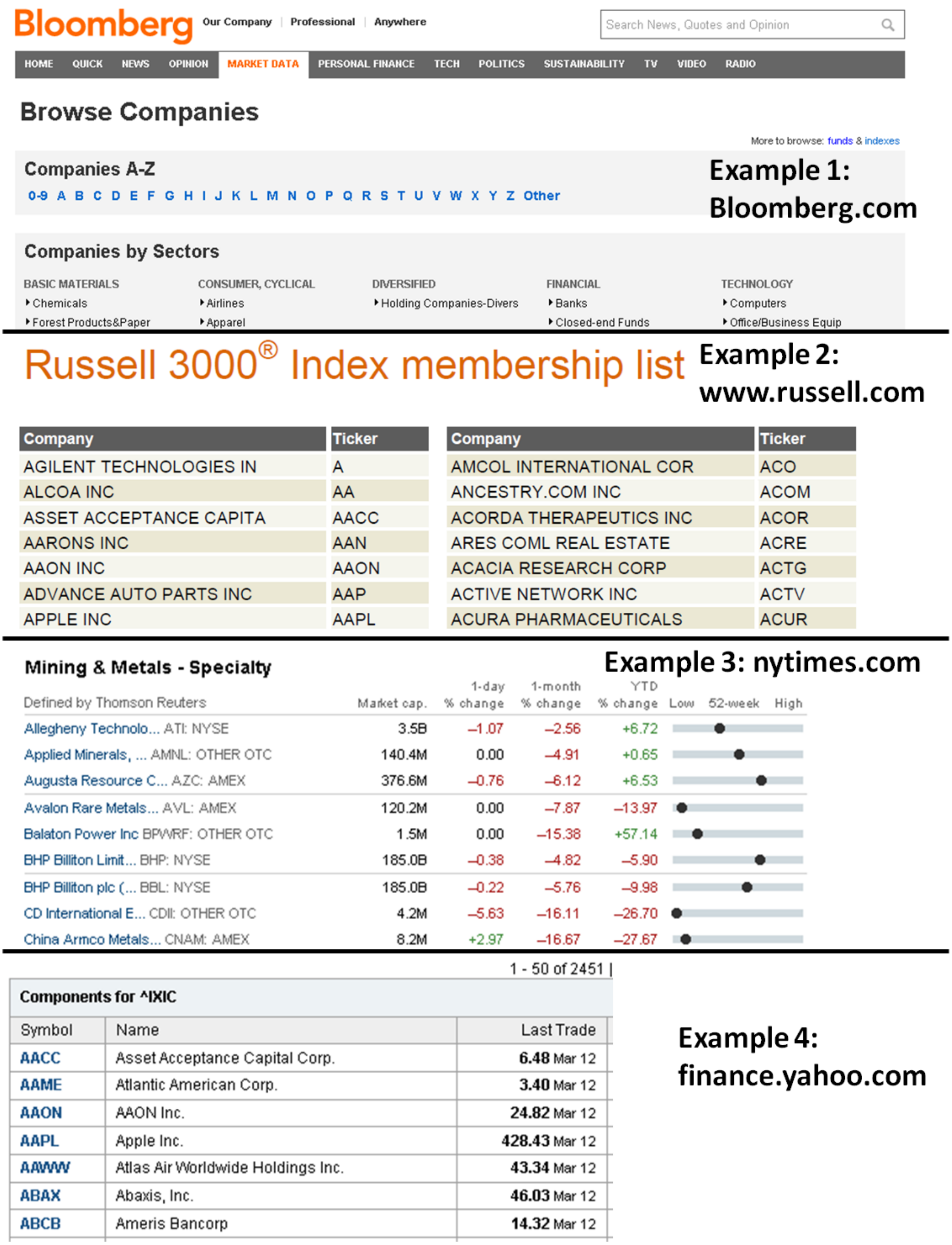
Solomon, D. H., E. F. Soltes, and D. Sosyura, 2013, "Winners in the spotlight: Media coverage of fund holdings as a driver of flows," *Journal of Financial Economics, forthcoming.*

Statman, M., S. Thorley, and K. Vorkink, 2006, "Investor overconfidence and trading volume," *Review of Financial Studies*, 19, 1531–1565.

Tetlock, P. C., 2011, "All the news that's fit to reprint: Do investors react to stale information?," *Review of Financial Studies*, 24, 1481–1512.

van Praag, C. M., and B. M. van Praag, 2008, "The benefits of being economics professor A (rather than Z)," *Economica*, 75, 782–796.

Figure 1: Motivating examples: Alphabetical ordering of stock names on popular finance websites



Bloomberg Our Company | Professional | Anywhere

Search News, Quotes and Opinion

HOME QUICK NEWS OPINION MARKET DATA PERSONAL FINANCE TECH POLITICS SUSTAINABILITY TV VIDEO RADIO

Browse Companies

More to browse: funds & indexes

Companies A-Z

0-9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z Other

Example 1: Bloomberg.com

Companies by Sectors

BASIC MATERIALS	CONSUMER, CYCLICAL	DIVERSIFIED	FINANCIAL	TECHNOLOGY
Chemicals	Airlines	Holding Companies-Divers	Banks	Computers
Forest Products&Paper	Apparel		Closed-end Funds	Office/Business Equip

Russell 3000® Index membership list

Example 2: www.russell.com

Company	Ticker	Company	Ticker
AGILENT TECHNOLOGIES IN	A	AMCOL INTERNATIONAL COR	ACO
ALCOA INC	AA	ANCESTRY.COM INC	ACOM
ASSET ACCEPTANCE CAPITA	AACC	ACORDA THERAPEUTICS INC	ACOR
AARONS INC	AAN	ARES COML REAL ESTATE	ACRE
AAON INC	AAON	ACACIA RESEARCH CORP	ACTG
ADVANCE AUTO PARTS INC	AAP	ACTIVE NETWORK INC	ACTV
APPLE INC	AAPL	ACURA PHARMACEUTICALS	ACUR

Mining & Metals - Specialty

Example 3: nytimes.com

Defined by Thomson Reuters	Market cap.	1-day % change	1-month % change	YTD % change	Low 52-week High
Allegheny Technolo... ATI: NYSE	3.5B	−1.07	−2.56	+6.72	
Applied Minerals, ... AMNL: OTHER OTC	140.4M	0.00	−4.91	+0.65	
Augusta Resource C... AZC: AMEX	376.6M	−0.76	−6.12	+6.53	
Avalon Rare Metals... AVL: AMEX	120.2M	0.00	−7.87	−13.97	
Balaton Power Inc BPWRF: OTHER OTC	1.5M	0.00	−15.38	+57.14	
BHP Billiton Limit... BHP: NYSE	185.0B	−0.38	−4.82	−5.90	
BHP Billiton plc (... BBL: NYSE	185.0B	−0.22	−5.76	−9.98	
CD International E... CDII: OTHER OTC	4.2M	−5.63	−16.11	−26.70	
China Armco Metals... CNAM: AMEX	8.2M	+2.97	−16.67	−27.67	

1 - 50 of 2451 |

Components for ^IXIC

Example 4: finance.yahoo.com

Symbol	Name	Last Trade
AACC	Asset Acceptance Capital Corp.	**6.48** Mar 12
AAME	Atlantic American Corp.	**3.40** Mar 12
AAON	AAON Inc.	**24.82** Mar 12
AAPL	Apple Inc.	**428.43** Mar 12
AAWW	Atlas Air Worldwide Holdings Inc.	**43.34** Mar 12
ABAX	Abaxis, Inc.	**46.03** Mar 12
ABCB	Ameris Bancorp	**14.32** Mar 12

Figure 2: Percentage difference in trading characteristics for alphabetically sorted firms

This figure compares trading activity (as measured by monthly stock-level turnover) and costs of trading (as measured by the monthly Amihud (2002) illiquidity ratio) for different groups of firms trading on the NYSE or AMEX. Firms are sorted alphabetically (in ascending order) by their firm name into five groups: First 5%, 5%-25%, 25%-50%, 50%-75%, and the last 25%, which serve as a benchmark. Findings display the percentage difference in trading activity and costs of trading for the first four groups relative to the fifth group. Values represent coefficients obtained from corresponding dummy variables in multivariate predictive panel regressions. Standard errors are double-clustered by firm and month. Additionally, year dummies are included in the regression. The sample period ranges from February 1995 to December 2009. Control variables are lagged (at least) one month and correspond to model 3 explained in section 2. The model controls for signed stock returns, book-to-market ratio, market capitalization, firm age, nominal share price, leverage ratio, beta, the number of analysts, analyst forecast dispersion, earnings surprises, earnings volatility, returns in t-12 and over t-7 to t-1, 52-week high, squared market capitalization, index membership (S&P 500, Dow Jones 30), idiosyncratic volatility, alpha, advertising expenditures, research and development expenditures, urban and rural regions, number of employees, industry classification, national and local press coverage (in total 45 newspapers), ten dummies for different types of 8-K filings, firm headquarter state, and company name fluency. Statistical significance at the ten, five and one-percent level is indicated by *, **, and ***, respectively.

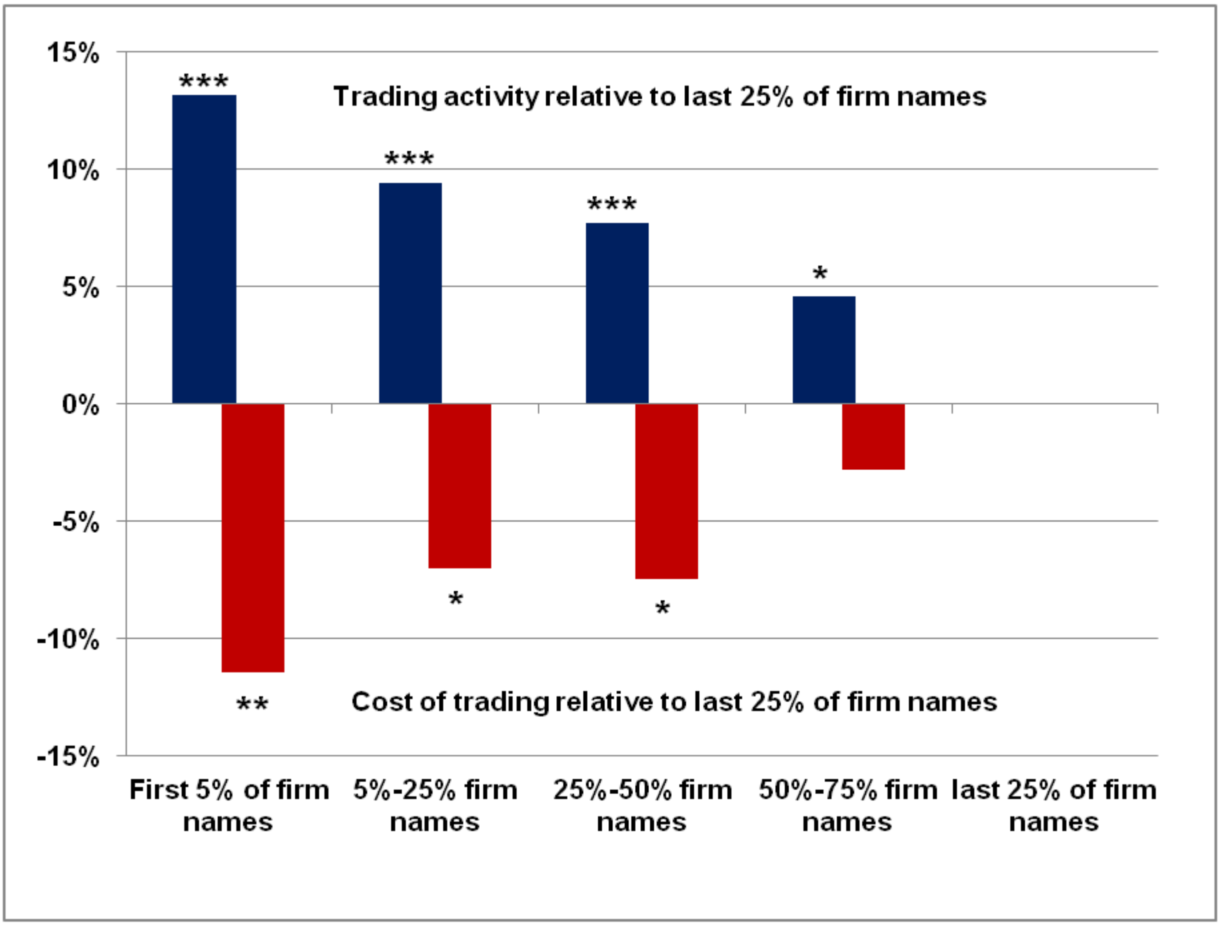


Table 1: Firm turnover and alphabetic bias: Predictive Fama/MacBeth regressions

This table summarizes the main results of nine predictive regressions which differ in firm universe (NYSE/AMEX, Nasdaq, all) and control variables/sample periods (models 1-3). The dependent variable is the natural logarithm of firm turnover in month t. In all regressions, the explanatory variable of interest is *position continuous*, defined as the relative position of a firm's name within the alphabetically sorted firm universe in the previous month t-1. Thus, *position continuous* takes on values between zero and one. All control variables are lagged at least by one month. If we run regressions across all exchanges, turnover for NYSE/AMEX firms is doubled, and a Nasdaq dummy is included. In panel A, model 1 controls for signed stock returns, book-to-market ratio, market capitalization, firm age, nominal share price, leverage ratio, and beta. In panel B, model 2 additionally controls for the number of analysts, analyst forecast dispersion, earnings surprises, earnings volatility, returns in t-12 and over t-7 to t-1, 52-week high, squared (logarithmized) market capitalization, index membership (S&P 500, Dow Jones 30, Nasdaq 100 (all where applicable)), idiosyncratic volatility, alpha, advertising expenditures, research and development expenditures, dummies for urban and rural regions, number of employees, and industry classification (49 Fama and French (1997) industries). Finally, model 3 additionally controls for national and local press coverage (in total 45 newspapers), all 8-K filings (ten different dummies which proxy for different firm events), firm headquarter state, and company name fluency as in Green and Jame (2013). % *predicted direction* denotes the fraction of monthly regressions which yield a negative coefficient for *position continuous*. T-statistics are based on heteroskedasticity- and autocorrelation-consistent (HAC) Newey and West (1987) standard errors. Statistical significance at the ten-, five- and one-percent level is indicated by *, **, and ***, respectively. The reported R^2 is calculated as the average of adjusted R^2.

Panel A: Model 1			
Firm universe	NYSE/AMEX	Nasdaq	All
Sample start	Jul-63	Jan-83	Jan-83
Sample end	Dec-11	Dec-11	Dec-11
Coefficient of *position continuous*	-0.1114***	-0.0547***	-0.0867***
t-stat of *position continuous*	-6.48	-2.78	-4.37
Fraction of neg. coefficients	87%	72%	83%
R^2	0.27	0.31	0.34
Average N	1,455	1,879	3,312
Panel B: Model 2			
Firm universe	NYSE/AMEX	Nasdaq	All
Sample start	Feb-80	Jan-83	Jan-83
Sample end	Dec-11	Dec-11	Dec-11
Coefficient of *position continuous*	-0.0914***	-0.0587***	-0.0804***
t-stat of *position continuous*	-5.44	-3.17	-7.64
Fraction of neg. coefficients	89%	75%	93%
R^2	0.43	0.43	0.45
Average N	1,349	1,486	2,820
Panel C: Model 3			
Firm universe	NYSE/AMEX	Nasdaq	All
Sample start	Feb-95	Feb-95	Feb-95
Sample end	Dec-09	Dec-09	Dec-09
Coefficient of *position continuous*	-0.1202***	-0.0686***	-0.0907***
t-stat of *position continuous*	-5.83	-3.66	-7.51
Fraction of neg. coefficients	96%	80%	97%
R^2	0.50	0.50	0.51
Average N	1,249	1,693	2,942

Table 2: (Il)liquidity and alphabetic bias: Predictive Fama/MacBeth regressions

This table summarizes the main results from nine predictive regressions which differ in firm universe (NYSE/AMEX, Nasdaq, all) and control variables/sample periods (models 1-3). The dependent variable is the logarithmized Amihud (2002) illiquidity ratio in month t. The ratio is computed as the monthly average of daily illiquidity measures, defined as the absolute daily stock return divided by daily trading volume in millions of dollars. In all regressions, the explanatory variable of interest is *position continuous*, defined as the relative position of a firm's name within the alphabetically sorted firm universe in the previous month t-1. Thus, *position continuous* takes on values between zero and one. All control variables are lagged at least by one month. Control variables in models 1 to 3 correspond to the ones used in table 1, where they are described in detail. % *predicted direction* denotes the fraction of monthly regressions which yield a positive coefficient for *position continuous*. T-statistics are based on heteroskedasticity- and autocorrelation-consistent (HAC) Newey and West (1987) standard errors. Statistical significance at the ten-, five- and one-percent level is indicated by *, **, and ***, respectively. The reported R^2 is calculated as the average of adjusted R^2.

Panel A: Model 1			
Firm universe	NYSE/AMEX	Nasdaq	All
Sample start	Jul-63	Jan-83	Jan-83
Sample end	Dec-11	Dec-11	Dec-11
Coefficient of *position continuous*	0.0879***	0.0756***	0.0916***
t-stat of *position continuous*	5.63	3.21	6.09
Fraction of pos. coefficients	83%	75%	90%
R^2	0.88	0.77	0.87
Average N	1,454	1,840	3,271
Panel B: Model 2			
Firm universe	NYSE/AMEX	Nasdaq	All
Sample start	Feb-80	Jan-83	Jan-83
Sample end	Dec-11	Dec-11	Dec-11
Coefficient of *position continuous*	0.0788***	0.0672***	0.0776***
t-stat of *position continuous*	5.07	2.90	7.41
Fraction of pos. coefficients	86%	77%	91%
R^2	0.92	0.81	0.89
Average N	1,349	1,482	2,816
Panel C: Model 3			
Firm universe	NYSE/AMEX	Nasdaq	All
Sample start	Feb-95	Feb-95	Feb-95
Sample end	Dec-09	Dec-09	Dec-09
Coefficient of *position continuous*	0.1093***	0.0883***	0.0952***
t-stat of *position continuous*	7.31	3.93	7.32
Fraction of pos. coefficients	93%	87%	94%
R^2	0.92	0.87	0.91
Average N	1,249	1,693	2,942

Table 3: Turnover, (il)liquidity and alphabetic bias: Alternative econometric approaches

This table summarizes the main results from thirthy predictive regressions which differ in the dependent variable (panel A: natural logarithm of firm turnover, panel B: logarithmized Amihud (2002) illiquidity ratio), in the firm universe (NYSE/AMEX, Nasdaq, all) and in the econometric approach (ID 1-5). In all regressions, control variables correspond to model 3 and are explained in detail in table 1. The sample period is February 1995 to December 2009. In all specifications, the explanatory variable of interest is *position continuous*, defined as the relative position of a firm's company name within the alphabetically sorted firm universe in the previous period t-1. Thus, *position continuous* takes on values between zero and one. In approach 1 (in both panels A and B), standard errors are double-clustered by firm and month. Additionally, year dummies are included in the regression. In approach 2, we rely on firm-fixed effects and on month-fixed effects. Approaches 3 to 5 rerun the analysis from the Fama and MacBeth (1973)-type regressions (see tables 1 and 2) as well as from approaches 1 and 2 at the yearly frequency. Statistical significance at the ten-, five- and one-percent level is indicated by *, **, and ***, respectively.

Panel A: Logarithmized turnover as dependent variable				
Approach ID and description		NYSE/AMEX	Nasdaq	All
1: Double-clustered standard errors,	Coefficient of *position continuous*	-0.1225***	-0.0812**	-0.1080***
monthly data	t-stat of *position continuous*	-3.46	-2.49	-4.21
2: Firm- and month-fixed effects,	Coefficient of *position continuous*	-0.0949***	-0.0342	-0.0918***
monthly data	t-stat of *position continuous*	-4.94	-1.39	-6.39
3: Fama/MacBeth approach,	Coefficient of *position continuous*	-0.1156***	-0.0726***	-0.0911***
yearly data	t-stat of *position continuous*	-9.12	-4.43	-10.45
4: Double-clustered standard errors,	Coefficient of *position continuous*	-0.1153***	-0.0915***	-0.1087***
yearly data	t-stat of *position continuous*	-3.30	-2.85	-4.45
5: Firm- and year-fixed effects,	Coefficient of *position continuous*	-0.0967**	-0.0298	-0.1073***
yearly data	t-stat of *position continuous*	-2.07	-0.57	-3.21

Panel B: Logarithmized Amihud illiquidity ratio as dependent variable				
Approach ID and description		NYSE/AMEX	Nasdaq	All
1: Doube-clustered standard errors,	Coefficient of *position continuous*	0.1115**	0.1052**	0.1072***
monthly data	t-stat of *position continuous*	2.24	2.49	3.15
2: Firm- and month-fixed effects,	Coefficient of *position continuous*	0.1647***	0.0781***	0.1427***
monthly data	t-stat of *position continuous*	7.90	3.16	8.71
3: Fama/MacBeth approach,	Coefficient of *position continuous*	0.1041***	0.0978***	0.096***
yearly data	t-stat of *position continuous*	12.12	4.52	8.10
4: Double-clustered standard errors,	Coefficient of *position continuous*	0.1035**	0.1194***	0.1103***
yearly data	t-stat of *position continuous*	2.22	2.89	3.40
5: Firm- and year-fixed effects,	Coefficient of *position continuous*	0.1388**	0.1092**	0.1614***
yearly data	t-stat of *position continuous*	2.45	2.01	3.71

Table 4: Turnover, (il)liquidity and alphabetic bias: Further evidence

This table summarizes the main results from six predictive panel regressions which differ in the dependent variable (Panel A: logarithmized monthly stock-level turnover, Panel B: logarithmized Amihud (2002) monthly illiquidity ratio), and the model specification (model 1-3 as described in detail in table 1). Displayed are the coefficients and t-statistics for four dummy variables, which indicate the position of a given firm name within the NYSE/Amex/Nasdaq universe. *Position dummy* 5 takes on a value of 1 if the company name is among the first 5% in a given month and zero otherwise. *Position dummy* 25 is 1 if the firm name is between percentiles 5 and 25. Finally, *position dummy* 50 (*position dummy* 75) is 1 if the name is between percentiles 25 and 50 (50 and 75). Consequently, findings are benchmarked against the firms near the end of the alphabet (percentile 75 and above). Standard errors are double-clustered by firm and month. Additionally, year dummies are included in the regression. Statistical significance at the ten-, five- and one-percent level is indicated by *, **, and ***, respectively.

Panel A: Logarithmized turnover as dependent variable			
Firm universe	All exchanges	All exchanges	All exchanges
Model specification	Model 1	Model 2	Model 3
Coefficient of position dummy 5	0.0894**	0.1069***	0.1074***
t-stat of position dummy 5	2.30	3.46	3.19
Coefficient of position dummy 25	0.0703***	0.0740***	0.0806***
t-stat of position dummy 25	3.01	3.87	3.61
Coefficient of position dummy 50	0.0413*	0.0516***	0.0626***
t-stat of position dummy 50	1.83	2.80	3.02
Coefficient of position dummy 75	0.0065	0.0121	0.0133
t-stat of position dummy 75	0.30	0.67	0.64
Panel B: Logarithmized illiquidity ratio as dependent variable			
Firm universe	All exchanges	All exchanges	All exchanges
Model specification	Model 1	Model 2	Model 3
Coefficient of position dummy 5	-0.1116***	-0.1156***	-0.1169***
t-stat of position dummy 5	-2.67	-3.12	-2.85
Coefficient of position dummy 25	-0.0530*	-0.0531**	-0.0623**
t-stat of position dummy 25	-1.74	-2.05	-2.07
Coefficient of position dummy 50	-0.0385	-0.0433*	-0.0642**
t-stat of position dummy 50	-1.39	-1.76	-2.32
Coefficient of position dummy 75	0.0220	0.0132	0.0101
t-stat of position dummy 75	0.78	0.54	0.36

Table 5: Breadth of ownership, firm valuation, and alphabetic bias: Panel regressions

This table summarizes the main results from panel regressions of breadth of ownership (panel A) and firm valuation (panel B) on measures of the relative position of a firm's name in alphabetical ordering and a number of controls. Breadth of ownership (firm valuation) is computed as the logarithmized number of shareholders (logarithmized market-to-book ratio) in a given year, averaged across all months in that year. In the case of market-to-book ratios, we exclude negative values as well as firms from the banking and insurance sector, and moreover winsorize the data at the 99th percentile in each year. Each panel reports findings from six regressions, which differ in the alphabetical ordering measure and the model specification used (model 1-3 as described in detail in table 1). In all regressions, we additionally include the stock's illiquidity ratio as a control variable. In all regressions, the firm universe consists of stocks trading at NYSE, Amex, or Nasdaq. Standard errors are double-clustered by firm and year. Additionally, year dummies are included in the regression. Statistical significance at the ten-, five- and one-percent level is indicated by *, **, and ***, respectively.

Panel A: Logarithmized number of shareholders as dependent variable			
Model specification	Model 1	Model 2	Model 3
N	95,129	82,032	45,663
Panel A1: Impact of position continuous			
Coefficient of position continuous	-0.0829**	-0.0726*	-0.0737
t-stat of position continuous	-2.10	-1.82	-1.41
Panel A2: Impact of position dummies			
Coefficient of position dummy 5	0.0802	0.0914	0.0852
t-stat of position dummy 5	1.45	1.55	1.16
Coefficient of position dummy 25	0.0852**	0.0663*	0.0701*
t-stat of position dummy 25	2.53	1.91	1.71
Coefficient of position dummy 50	0.0347	0.0252	0.0515
t-stat of position dummy 50	1.09	0.78	1.21
Coefficient of position dummy 75	0.0617*	0.0404	0.0341
t-stat of position dummy 75	1.89	1.24	0.78
Panel B: Logarithmized market-to-book-ratio as dependent variable			
Model specification	Model 1	Model 2	Model 3
N	93,052	80,216	44,550
Panel B1: Impact of position continuous			
Coefficient of position continuous	-0.0462**	-0.0104	-0.0279
t-stat of position continuous	-2.01	-0.55	-1.35
Panel B2: Impact of position dummies			
Coefficient of position dummy 5	0.0905***	0.0202	0.0578**
t-stat of position dummy 5	2.78	0.81	2.01
Coefficient of position dummy 25	0.0086	-0.0011	0.0068
t-stat of position dummy 25	0.42	-0.07	0.38
Coefficient of position dummy 50	0.0331*	-0.0042	0.0045
t-stat of position dummy 50	1.68	-0.30	0.26
Coefficient of position dummy 75	0.0009	-0.0101	-0.0087
t-stat of position dummy 75	0.05	-0.76	-0.52

Table 6: Investor sophistication and alphabetic bias

This table summarizes main results from pooled predictive panel regressions of stock turnover and illiquidity (panel A) or stock turnover and breadth of ownership (panel B) on *position continuous* (the coefficient of interest) and control variables (see table 1 for details). Panel A shows subsample regressions based on model 3 and NYSE/AMEX/Nasdaq stocks. The sample period ranges from February 1995 to December 2009. Small firms (large firms) are firms with a below (above) median market capitalization in the previous month. Lottery firms (non-lottery firms) are firms with above (below) median idiosyncratic volatility, above (below) median idiosyncratic skewness, and below (above) median nominal share price. Firms with high (low) media coverage are firms with above (below) median residual media coverage, defined as the residual from monthly cross-sectional regressions of the logarithmized number of firm-specific articles in the New York Times, the USA Today, the Wall Street Journal, and the Washington Post on logarithmized lagged market capitalization. High (low) institutional ownership refers to firms with above (below) median institutional ownership lagged by one quarter. Panel B compares the behavior of retail investors from a large discount broker with the overall market behavior. Regressions are either based on model 1 or 2. The sample period ranges from January 1991 to November 1996. Retail turnover is computed as the logarithmized ratio of the monthly number of shares traded by individual investors and the number of shares outstanding. We consider all stocks which meet the data requirements of model 1 or 2, respectively. Retail breadth of ownership for a given stock in a given month is ln (1+number of individual investors who own the stock). Both retail turnover and retail breadth of ownership are winsorized each month at the 95th percentile. In both panels, standard errors are double-clustered by firm and month. Additionally, year dummies are included in the regression. Statistical significance at the ten-, five- and one-percent level is indicated by *, **, and ***, respectively.

Panel A: Stock subsamples					
Logarithmized turnover as dependent variable			Logarithmized illiquidity ratio as dependent variable		
	Coefficient	t-stat		Coefficient	t-stat
Small stocks (N=263,192)	-0.1446***	-4.03	Small stocks (N=263,136)	0.1835***	3.96
Large stocks (N=263,314)	-0.0684**	-2.48	Large stocks (N=263,314)	0.0374	0.91
Difference	-0.0762*	-1.78	Difference	0.1461**	2.47
Lottery stocks (N=125,116)	-0.1122***	-2.79	Lottery stocks (N=125,099)	0.1427***	2.78
Non-lottery stocks (N=122,015)	-0.0923***	-3.14	Non-lottery stocks (N=122,007)	0.0909*	1.94
Difference	-0.0199	-0.39	Difference	0.0518	0.76
High media coverage (N=263,286)	-0.1321***	-3.95	High media coverage (N=263,233)	0.1502***	3.63
Low media coverage (N=263,330)	-0.0775***	-2.68	Low media coverage (N=263,217)	0.0578	1.42
Difference	-0.0546	-1.44	Difference	0.0924*	1.90
Low institutional ownership (N=263,184)	-0.1714***	-4.45	Low institutional ownership (N=263,133)	0.1914***	4.06
High institutional ownership (N=263,322)	-0.0266	-1.10	High institutional ownership (N=263,317)	0.0146	0.35
Difference	-0.1448***	-3.30	Difference	0.1768***	2.97
Panel B: Retail investor behavior vs. market behavior					
Turnover (Model 1)	Retail	Market	Turnover (Model 2)	Retail	Market
N	248,644	248,644	N	203,835	203,835
R^2	0.16	0.25	R^2	0.23	0.37
Coefficient of *position continuous*	-0.1649*	-0.0797*	Coefficient of *position continuous*	-0.1395	-0.0872**
t-stat of *position continuous*	-1.71	-1.80	t-stat of *position continuous*	-1.50	-2.44
Breadth of ownership (Model 1)	Retail	Market	Breadth of ownership (Model 2)	Retail	Market
N	245,686	20,954	N	203,337	17,535
R^2	0.60	0.59	R^2	0.67	0.67
Coefficient of *position continuous*	-0.0774**	-0.0614	Coefficient of *position continuous*	-0.0583	-0.0462
t-stat of *position continuous*	-2.01	-1.45	t-stat of *position continuous*	-1.56	-1.10

Table 7: Mutual fund flows and alphabetic bias

This table summarizes the main results from predictive regressions of monthly mutual funds flows (winsorized each month at the 5% and 95% level) on measures of the relative position of fund names within the alphabetically sorted fund universe as well as on a number of control variables. The sample period ranges from January 1992 to December 2012. We focus on open-end U.S. equity mutual funds with a domestic investment focus. In specifications 1 and 2, we use all funds. Models 3 and 4 (5 and 6) only consider funds with total net assets below or equal to (above) the median-sized fund in a given month. *Position continuous* as well as the four *Position dummies* are defined analogously to the analysis in the stock market (see e.g. table 1 and 4). *Fund size* is the natural logarithm of one-month lagged total net assets of the fund (in million USD). *Fund age* is ln(age in months). *Fund expense ratio* and *fund turnover* are lagged by one year and logarithmized. *Length of fund name* is the number of letters of the fund name, after dropping share-class information and incorporation terms where applicable. *Fund return* is the cumulative return (net of fees) over the previous twelve months. *Fund risk* is the standard deviation of the previous twelve monthly return observations. *Fund performance rank* indicates the relative performance of the fund in its market segment (one of 19 styles identified by the CRSP style code) in a given month. At the family-level, *total net assets* and *no. of funds* are lagged by one month and logarithmized. *Top 3 fund families* is a dummy which takes on a value of one (zero) if the fund belongs to one of the three largest fund families based on one-month lagged total net assets. *Style growth* is the current growth rate of the fund's market segment (one of 19 styles) due to aggregated fund flows. All regressions contain dummies for each *style − year observation*. Standard errors are double-clustered by fund and month. Statistical significance at the ten-, five- and one-percent level is indicated by *, **, and ***, respectively.

Fund universe	All funds		Small funds		Large funds	
Model specification	(1)	(2)	(3)	(4)	(5)	(6)
Position continuous	-0.135**		-0.207***		-0.0248	
	(-2.44)		(-2.82)		(-0.33)	
Position dummy 5		0.226***		0.315***		0.0997
		(3.03)		(3.28)		(1.02)
Position dummy 25		0.0173		0.0904		-0.0918
		(0.36)		(1.44)		(-1.40)
Position dummy 50		0.0407		0.0378		0.0156
		(0.88)		(0.60)		(0.25)
Position dummy 75		-0.141***		-0.0729		-0.225***
		(-3.04)		(-1.24)		(-3.56)
Fund size	0.237***	0.238***	0.202***	0.200***	0.237***	0.238***
	(16.77)	(16.80)	(8.67)	(8.55)	(9.19)	(9.22)
Fund age	-0.932***	-0.934***	-0.961***	-0.962***	-0.902***	-0.906***
	(-24.97)	(-25.03)	(-21.43)	(-21.48)	(-19.46)	(-19.58)
Fund expense ratio	0.145**	0.147***	0.108*	0.109*	0.196**	0.206**
	(2.57)	(2.64)	(1.75)	(1.77)	(2.15)	(2.27)
Fund turnover rate	-0.0145	-0.0133	-0.0166	-0.0170	-0.0134	-0.00909
	(-0.71)	(-0.65)	(-0.68)	(-0.69)	(-0.47)	(-0.32)
No. share classes	0.0085	0.0353	0.0910	0.110	-0.241*	-0.228*
	(0.08)	(0.34)	(0.69)	(0.83)	(-1.76)	(-1.66)
Length of fund name	-0.0019	-0.0020	-0.0033*	-0.0036**	-0.0005	-0.0007
	(-1.40)	(-1.48)	(-1.83)	(-1.96)	(-0.29)	(-0.40)
Fund return	3.385***	3.382***	3.204***	3.202***	3.517***	3.508***
	(11.23)	(11.23)	(10.42)	(10.42)	(11.10)	(11.09)
Fund risk	0.269	0.297	0.298	0.326	0.365	0.388
	(0.16)	(0.18)	(0.17)	(0.19)	(0.18)	(0.19)
Fund performance rank	0.197	0.203	0.294*	0.299*	0.116	0.130
	(1.34)	(1.38)	(1.67)	(1.70)	(0.63)	(0.71)
$(\text{Fund performance rank})^2$	0.626***	0.619***	0.531***	0.526***	0.686***	0.671***
	(4.60)	(4.56)	(3.20)	(3.17)	(3.98)	(3.90)
Fund family: Total net assets	0.184**	0.186**	0.140	0.144	0.399	0.449
	(2.09)	(2.10)	(1.33)	(1.37)	(1.41)	(1.59)
Fund family: No. of funds	-0.186***	-0.189***	-0.169***	-0.169***	-0.221***	-0.233***
	(-6.58)	(-6.64)	(-4.23)	(-4.21)	(-4.93)	(-5.15)
Top 3 fund families	0.384***	0.371***	0.674***	0.695***	0.305***	0.260***
	(5.28)	(4.99)	(5.03)	(5.11)	(3.62)	(2.95)
Style growth	78.33***	78.34***	70.20***	70.22***	87.92***	87.94***
	(27.32)	(27.33)	(22.13)	(22.15)	(26.00)	(26.03)
Constant	2.225***	2.226***	2.412***	2.323***	3.349***	3.544***
	(4.07)	(4.05)	(4.12)	(3.95)	(5.56)	(5.91)
Style-year fixed effects	yes	yes	yes	yes	yes	yes
Observations	374,566	374,566	187,354	187,354	187,212	187,212
R-squared	0.172	0.172	0.141	0.141	0.220	0.221