Review of Finance, 2016, 693–723
doi: 10.1093/rof/rfv060
Advance Access Publication Date: 14 December 2015

Alphabetic Bias, Investor Recognition, and Trading Behavior*

Heiko Jacobs and Alexander Hillert

University of Mannheim

Abstract

Extensive research has revealed that alphabetical name ordering tends to provide an advantage to those positioned in the beginning of an alphabetical listing. This article is the first to explore the implications of this alphabetic bias in financial markets. We find that US stocks that appear near the top of an alphabetical listing have about 5–15% higher trading activity and liquidity than stocks that appear toward the bottom. The magnitude of these results is negatively related to firm visibility and investor sophistication. International evidence and fund flows further indicate that ordering effects can affect trading activity and liquidity.

JEL classification: G02, G12, G14

1. Introduction

Being on or near the top of a list is in many ways important in academia. For example, Brogaard, Engelberg, and Parsons (2014) show that articles placed first, second, and third in an issue of top finance and economic journals receive about 50%, 26%, and 17% more

* Parts of this research project were conducted while the first author visited the School of Banking and Finance at the University of New South Wales in 2013. We would like to thank (in alphabetical order . . .) Tim Adam, Daniel Andrei, James Choi (AFA discussant), Russell Jame, Markku Kaustia (Helsinki discussant), Alexander Klos, Alok Kumar, Ron Masulis, Ernst Maug, Ralitsa Petkova (EFA discussant), Erik Theissen, and Martin Weber for valuable comments. Moreover, we express our thanks to an anonymous referee whose insightful comments significantly improved the article. In addition, we thank seminar participants at the 3rd European Retail Investment Conference (ERIC) 2015, at the 75th Annual Meeting of the American Finance Association (AFA) 2015, at the 41st Annual Meeting of the European Finance Association (EFA) 2014, at the 76th Annual Meeting of the German Academic Association for Business Research (VHB) 2014, at the Third Helsinki Finance Summit on Investor Behavior 2013, at the 20th Annual Meeting of the German Finance Association (DGF) 2013, at the University of Mainz (2015), at the Humboldt University of Berlin/ESMT (2014), at the University of New South Wales (2013), and at the University of Mannheim (2013, 2014). We are grateful for financial support from the Julius Paul Stiegler Memorial Foundation. An earlier draft of this article was circulated under the title “The Power of Primacy: Alphabetic Bias, Investor Recognition, and Market Outcomes”.

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 the editors' private information about article quality or the readers' limited attention. Other evidence, however, points more directly to the latter explanation,[1] and also indicates that rank effects are often strongly related to the widespread convention of alphabetization. For instance, Einav and Yariv (2006) and van Praag and van Praag (2008) demonstrate that, due to the increased visibility of first-named authorship in co-authored publications in economics, researchers with last name initials closer to the beginning of the alphabet benefit by various measures of academic success. These measures include a better chance of receiving tenure at a top ten economics department or a prestigious award.

There are other examples for this form of alphabetic bias in academia. As editors of a well-established medical imaging journal rely on an alphabetically ordered list of potential referees, reviewers whose last name starts with an A receive almost twice as many review invitations as their colleagues whose last name starts with a letter toward the end of the alphabet (Richardson, 2008). Similar mechanisms apply to positions as visiting academics (Merritt, 1999) and to the amount of attention directed toward university alumni (Meer and Rosen, 2011).

As the convention of alphabetization is a fact of everyday life and as browsing from the top of lists is natural human behavior, alphabetic bias is not limited to academia. For instance, the bias 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 in political elections (e.g., Ho and Imai, 2008). Collectively, this leads "The Economist" (2001, p. 13) 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".

To the best of our knowledge, we are the first to deal with the role of alphabetization in the natural setting of the stock market. There is a widespread convention of listing stocks alphabetically in newspapers, online information sources, indices, data bases, broker statements, and watch lists. Given the thousands of stocks available, investors encounter a search problem, making ordering effects in financial decisions likely to occur. Indeed, both the alphabetical distribution of firm names and anecdotal evidence indicate that at least some firms care about being near the top of alphabetically ordered lists.

Within the context of stock markets, we explore the implications of alphabetic bias for the cross-section of stock-level trading activity and liquidity. Our choice is motivated both by the straightforward behavioral link and by the economic importance of these financial market outcomes.[2]

1 For instance, despite randomized list placement, the NBER working paper ranked first in a widely distributed weekly email announcement generates 33% more hits, 30% more downloads, and 27% more citations (Feenberg *et al.*, 2015).

2 Trading variables are arguably the most plausible outcome variables. Our clear focus allows us to comprehensively investigate the robustness, drivers, and implications of the findings in different markets and among different groups of investors in a structured way. Nevertheless, alphabetic bias may leave other traces in financial markets as well. For instance, untabulated results indicate that the increased demand for stocks that appear near the top of an alphabetical listing may result in higher market-to-book ratios. However, this effect seems to be less robust, economically smaller,

There are several (partly interrelated) mechanisms through which a higher alphabetical ranking may lead to greater trading volume and liquidity. First, investors can only trade stocks that are part of their consideration set, as formalized, for example, in the model of Merton (1987). Consequently, more visible firms have a higher chance of being selected for trading and also tend to be more liquid (e.g., Grinblatt and Keloharju, 2001a; Loughran and Schultz, 2005; Barber and Odean, 2008). Being placed near the top of an alphabetically ordered stock list is likely to boost visibility, thereby potentially stimulating both trading activity and liquidity. Second, this increased visibility could translate into a feeling of familiarity, which may lead to an increase in perceived knowledge and competence of evaluating a stock's prospects (e.g., Heath and Tversky, 1991). This familiarity can induce both higher trading activity (e.g., Huberman, 2001) and liquidity (e.g., Grullon, Kanatas, and Weston, 2004). Third, even in the absence of permanent visibility or familiarity effects, alphabetic bias could play a role in the decision in which firm to invest, thereby affecting stock-level trading activity. Market participants might run out of time, capital, interest, or face other constraints before they reach the bottom of a list with alphabetically ordered investment alternatives. Even if investors worked through a number of firms from the top to the bottom of, for instance, a typical stock index, the well-documented primacy effect (e.g., Carney and Banaji, 2012) suggests that they may prefer one of the first listed stocks. Finally, some investors might (perhaps unconsciously) associate the first stock on an alphabetically ordered list with superior quality (e.g., Fedenia and Hirschey, 2009; Ang, Chua, and Jiang, 2010).

Trading activity and liquidity are intrinsic features of financial markets, and they are economically relevant variables. As discussed in detail in the literature survey of Bond, Edmans, and Goldstein (2012), liquidity can affect real outcomes through the learning channel or the incentive channel. More specifically, liquidity encourages informed trading, which increases price informativeness. Greater informativeness in turn can allow the manager and other market participants to learn from prices. Separately, the greater price informativeness that arises from higher liquidity means that stock prices reflect firm value to a greater extent. Thus, the manager, whose compensation is tied to the stock price, has greater incentives to improve firm value. In sum, we explore the role of alphabetic bias in a novel and important setting in which ordering effects could have real consequences.

Our findings, which are based on predictive panel regressions and a sample period ranging from January 1983 to December 2011, show that US firms positioned early in an alphabetically ordered list indeed have higher trading activity and liquidity. For instance, NYSE/Amex/Nasdaq firms in the top 5% of the alphabetically ordered stock universe have about 12% higher monthly turnover and 13% lower monthly Amihud (2002) illiquidity ratios than firms in the lowest 25%. These differences are statistically significant at the 1% level. In deriving these findings, we control for about twenty firm characteristics known to be related to trading behavior (e.g., market capitalization, share price, industry membership, analyst coverage, past returns, and media coverage).

Cross-sectional drivers of the effect are in line with both theoretical predictions (e.g., Merton, 1987) and stylized facts (e.g., Barber and Odean, 2008). More specifically, alphabetic bias is negatively related to firm visibility, as proxied for by firm size, by analyst

and statistically less significant than the findings obtained for trading variables. In line with related work (e.g., Hong, Kubik, and Stein, 2008; Green and Jame, 2013), we find no significant relation between a stock's alphabetical ranking and its returns.

coverage orthogonalized with respect to firm size, by industry market share, and by a composite measure which condenses the information from the three individual variables. In addition, mutual fund trading estimated from quarterly holdings, actual trading data from a large discount brokerage as well as subsample tests based on the fraction of institutional ownership suggest that alphabetic bias is negatively related to investor sophistication.

The impact of a firm's alphabetic position on stock-level trading activity and liquidity is roughly comparable across the disjunct samples of NYSE/Amex stocks, Nasdaq stocks, and stocks in major international markets. In addition, our findings are robust with respect to plausible modifications to the explanatory variable of interest (e.g., quantifying a firm's alphabetic position based on its ticker symbol) and to the econometric approach (e.g., using Fama and MacBeth (1973)-type regressions). With respect to changes of the dependent variable, our findings indicate that alphabetic bias may be more important for the cost-per-volume channel of liquidity (as, for instance, proxied for by the Amihud (2002) illiquidity ratio) than for the cost-per-price channel (as, for instance, proxied for by the two spread-based measures of Corwin and Schultz (2012) and Chung and Zhang (2014)). However, among otherwise less visible stocks, each of our alternative proxies for trading activity and liquidity (the aforementioned variables, dollar volume, share volume, number of trades, and the fraction of zero return days) indicates an economically meaningful and statistically significant (*t*-statistics 2.21–4.72) impact of alphabetic bias.

It seems unlikely that our findings are driven by an omitted variable that would need to be correlated with both alphabetic ranking and trading activity. First, empirically, a firm's alphabetic position appears to be largely unrelated to those firm characteristics that have previously been linked to cross-sectional differences in turnover and liquidity (e.g., share price or industry membership). Second, this variable would additionally need to be able to explain why the effect of alphabetic bias is much stronger among less visible firms. Third, insights from name changes are similar in economic magnitude although not statistically significant. Fourth, we do not find alphabetic bias in Japan, where firms are sorted based on a numerical stock code unrelated to alphabetic ordering. In line with the idea of rank effects, it is instead the relative position within the numerical code system that is predictive of trading activity in Japan.

As a further out-of-sample test, we turn to the mutual fund market. The mutual fund business is a trillion dollar market, making the cross-sectional distribution of fees earned by fund companies economically important. Further economic importance is gained from the fact that flows into mutual funds can affect the prices of stocks held by the funds (e.g., Coval and Stafford, 2007; Edmans, Goldstein, and Jiang, 2012). Thus, mutual fund flows can affect real economic activity, making it another important testing ground for the impact of alphabetic bias.

The mechanisms through which alphabetical ranking may affect trading behavior in the mutual fund market are largely comparable to the mechanisms through which alphabetical ranking may affect trading behavior in the stock market. For instance, with respect to the role of visibility, Sirri and Tuffano (1998, p. 1607) test and verify the hypothesis that "consumers would purchase those funds that are easier or less costly for them to identify." Empirical work (e.g., Barber, Odean, and Zheng, 2005; Cooper, Gulen, and Rau, 2005a) provides further support for this idea of cognitive overload in the mutual fund investment decision. As investors face multiple investment alternatives, funds with names that appear earlier in an alphabetical listing may more likely be part of investors' choice set. Combined with the idea of a persistent stream of net new money flows—French (2008) documents that open-end funds accounted for 10.5% of US equity in 1990, but for 32.4% in 2007—this line of reasoning gives rise to the conjecture that funds near the top of an alphabetical listing could grow faster.

Empirically, we follow the standard in the literature (e.g., Sirri and Tufano, 1998; Green and Jame, 2013) and analyze the cross-section of monthly flows between 1992 and 2012. We find that the effect of alphabetic fund name ordering is concentrated in small funds for which search costs tend to be higher. In this subgroup, funds in the first 5% of an alphabetical listing have about 3–6% more net inflows per annum. In deriving this finding, we control for fund characteristics (such as past performance), fund family characteristics (such as number of funds), and style characteristics (such as style growth).

Our study makes several contributions to the literature. First, it adds to the emerging research on name effects in financial markets. This literature shows that factors unrelated to fundamentals can influence investor behavior and market outcomes. In a clinical study, Rashes (2001) uncovers excess comovement between two stocks with similar ticker symbols and attributes this to confusion among predominantly small investors. Examples for name effects are also presented in Cooper, Dimitrov, and Rau (2001) and Cooper et al. (2005b): During (after) the dot.com bubble, firms that simply changed their name to (away from) "dot.com names" generated positive abnormal post-announcement returns. Similarly, cosmetic name changes among mutual funds yield positive abnormal inflows without delivering better performance (Cooper, Gulen, and Rau, 2005a). Green and Jame (2013) show that firms with short, easy to pronounce, memorable names attract more investors, which translates into higher stock-level turnover, lower transaction price impact, and higher firm value. Similarly, funds with more fluent names attract higher inflows. Our contribution is to show that even the ordering of names, and not the meaning of the names themselves, matters for trading activity and liquidity.

Relying on US stocks and a sample period from 1985 to 2012, a subsequent paper by Itzkowitz, Itzkowitz, and Rothbort (2015) confirms our finding that stocks that appear higher in an alphabetical listing have higher trading activity. The authors argue that this translates into higher valuations, and that the magnitude of the findings increases with the introduction of online trading and online access to stock lists. They mainly attribute their results to investors' being prone to search satisficing bias and status quo bias in the stock selection process. In contrast to our study, Itzkowitz, Itzkowitz, and Rothbort (2015) do not analyze the alphabetic distribution of firm names, name changes, international stock markets, portfolio-level trading decisions, and mutual funds.

Second, our findings add to a more comprehensive understanding of the puzzling cross-sectional differences in stock-level trading (see, e.g., the discussions in Chordia, Huh, and Subrahmanyam (2007) or Lo and Wang (2000)). We show that a previously neglected and seemingly minor detail, essentially the first letter of a company name, has a strong impact on investors' trading decisions. Thus, seemingly innocuous choices by firms can affect how they are traded in financial markets.

Third, our study adds to the 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 consider all alternatives simultaneously. Both theory (e.g., Barberis and Shleifer, 2003) and empirical evidence (e.g., Barberis, Shleifer, and Wurgler, 2005) suggest that investors instead may resort to complexity-reducing heuristics. Consistent with the idea of salience induced by ordering (Hartzmark, 2015), our contribution is to comprehensively explore a largely neglected channel through which attention allocation of different groups of market participants can be assessed.

Fourth, from a methodological point of view, our insights suggest that alphabetical ordering may be a plausible instrument for trading variables. More specifically, in a setting in which trading activity or liquidity is a key explanatory variable but in which endogeneity is a concern, (alphabetic) ordering may provide a useful source of exogenous variation in these qualities. This idea may be usable in future research.

2. Empirical Approach and Baseline Findings

2.1 Dependent Variables

The dependent variables are stock-level measures of trading activity and liquidity. As a baseline proxy for trading activity, we rely on stock turnover, defined as the number of shares traded divided by the number of shares outstanding. Turnover is potentially the most widely used proxy for trading activity in the literature (e.g., Campbell, Grossman, and Wang, 1993). If the two-fund separation theorem holds, then turnover (in contrast to, e.g., dollar volume or share volume) has to be the same for all risky assets. This theoretical consideration offers a benchmark against which we can evaluate the cross-section of trading activity. Thus, judging from the "relation between volume and equilibrium models of asset markets (. . .), turnover yields the sharpest empirical implications and is the most natural measure" (Lo and Wang, 2000, p. 12). As a baseline proxy for liquidity, we rely on the Amihud (2002) illiquidity ratio, defined as the monthly average of the ratio of the absolute daily stock return and daily trading volume in millions of dollars. This approach is widely used, easy to compute, intuitive to interpret, and, according to the horse race of liquidity measures performed in Goyenko, Holden, and Trzcinka (2009, p. 154) "a good proxy for price impact."

In our baseline analysis, we use log transformations of both dependent variables for the following reasons.[3] First, by construction, both measures cannot be negative, which results in extremely right-skewed distributions. Log transformations help to mitigate the impact of outliers. On a related note, Lo and Wang (2000, p. 282) document that the cross-section of log-turnover has desirable statistical properties in that it is "similar over time up to a location parameter". Second, previous empirical work (e.g., Grullon, Kanatas, and Weston, 2004; Keloharju, Knpfer, and Linnainmaa, 2012; Green and Jame, 2013) suggests that cross-sectional differences in investor recognition have less explanatory power for absolute differences (additive) in trading variables, but may be more informative about relative differences (multiplicative). Third, and relatedly, quantifying the relative impact of alphabetic bias by using a logged dependent variable better captures the type of relationships that we are primarily interested in. More specifically, the approach enables us to easily draw meaningful comparisons of the effect size between subsamples of stocks (e.g., visible versus less visible) or between investor groups (e.g., fund managers versus overall market) that

3 To avoid the problem of zero turnover and to make the distribution of turnover closer to normal, we follow Llorente *et al.* (2002) in adding a constant of 0.00000255 before taking logs. Inferences do not depend on this choice. For instance, findings are similar if we do not add a constant at all. However, using an arbitrary constant which is small relative to the typical level of turnover yields desirable distributional properties for an OLS regression.

materially differ in their unconditional trading characteristics. Nevertheless, Section 4 shows that inferences remain unchanged if we measure the dependent variables in absolute units.

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), we run our analysis separately for NYSE/Amex and Nasdaq stocks.[4] Similar patterns across these disjunct subsamples would suggest that our findings represent a more generalized phenomenon. To additionally estimate an overall effect across exchanges, we follow the method suggested in, for example, Loughran and Schultz (2005) by doubling trading volume for NYSE and Amex stocks, pooling them with Nasdaq stocks, and adding a Nasdaq dummy variable to the regression.

Econometrically, we rely on the framework proposed in Chordia, Huh, and Subrahmanyam (2007). The approach involves running predictive regressions of a measure of trading activity for stock i in month $t+1$ ($\text{Trading}_{i,t+1}$) on n lagged explanatory firm characteristics, which we define in the following two sections:

$$\text{Trading}_{i,t+1} = \gamma_{0,t} + \sum_{k=1}^{n} \gamma_{k,t}\,\text{Firm Characteristic}_{k,i,t} + \epsilon_{i,t+1}. \tag{1}$$

2.2 Control Variables

All of the following variables have been proposed as determinants of the cross-section of stock-level trading activity or liquidity in related literature. Details about their construction and about the data sources are given in Table 1 of the Online Appendix. Unless noted otherwise, these controls are used in all our stock market tests.

We start with all explanatory variables used in the full model specification of Chordia, Huh, and Subrahmanyam (2007). We include two variables defined as the stock return in the previous month if positive (negative) and zero otherwise. Several behavioral theories and stylized empirical facts suggest that past returns could influence trading activity. First, many investors do not hold the market portfolio (e.g., Goetzmann and Kumar, 2008), but instead may use heuristics such as the $1/N$ rule (e.g., Benartzi and Thaler, 2007). As a consequence, changes in asset valuations could lead to the need to rebalance and thus motivate trading. Second, past extreme returns may lead to attention-driven trading (e.g., Barber and Odean, 2008). Third, and as jointly formalized in Hong and Stein (1999), both positive-feedback trading and slow information diffusion can give rise to returns having predictive power for turnover. The findings in Barber and Odean (2008) or Grinblatt and Keloharju (2001b) indicate that the impact of positive and negative returns may be asymmetric, even though the proposed drivers make partly contrasting predictions. For instance, a stronger role of negative returns could be partly caused by tax-loss selling, a stronger role of positive returns is suggested by short-selling constraints (in combination with, for instance, positive-feedback trading) or by the disposition effect.

We employ the book-to-market ratio to account for cross-sectional differences in investor attention or dispersion of beliefs, both of which may explain why growth stocks tend

4 The NYSE is primarily an auction market, where a transaction of ten shares results in a total trading volume of ten shares. In contrast, the Nasdaq is a dealer market where buy and sell orders are executed against dealer quotes. As a consequence, buying and selling are viewed as different actions so that a transaction of ten shares results in a total trading volume of twenty shares.

to have higher turnover than value stocks (e.g., Hong and Stein, 2007). 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 other factors that may affect turnover or liquidity (e.g., Merton, 1987; Choi, Getmansky, and Tookes, 2009; Kumar, 2009). While these three variables tend to be substantially correlated, our findings are robust to specifications in which we only include one or two of them. We use log(1+ number of analysts providing fiscal year one earnings estimates in $t-1$) as a proxy for information diffusion, the mass of informed agents, and the firm's visibility (e.g., Hong, Lim, and Stein, 2000). We use analyst forecast dispersion and a firm's leverage as proxies for differences of opinion that may motivate trading (e.g., Diether, Malloy, and Scherbina, 2002). As in Chordia, Huh, and Subrahmanyam (2007), estimation uncertainty about fundamental values is quantified by lagged measures of beta, earnings surprises, and earnings volatility. Stronger estimation uncertainty may translate into higher trading activity due to a stronger need for rebalancing trades or as learning-induced trading may be more likely.

We augment the model (for month $t+1$) with additional explanatory variables that are not proposed in Chordia, Huh, and Subrahmanyam (2007), but that have been studied in related work. More specifically, we employ the return 12 months ago ($t-11$), the cumulative return over $t-7$ to $t-1$, and a 52-week high dummy to account for the increased visibility of these stocks, as discussed in, for example, Heston and Sadka (2008) or Seasholes and Wu (2007). We include idiosyncratic volatility, computed as the standard deviation of the residuals of rolling regressions of a firm's excess return on a standard four-factor model, as a proxy for firm-level disagreement, attention, and popularity among retail investors (e.g., Brandt *et al.*, 2010). The alpha from this regression is also included as it can be related to heterogeneous information (e.g., Lo and Wang, 2000).

To control for index effects, we include dummy variables for S&P 500 membership and Dow Jones 30 membership (e.g., Chen, Noronha, and Singal, 2004). Being included in an index could go along with increased visibility or increased trading by index funds or benchmark-oriented mutual funds. Grullon, Kanatas, and Weston (2004) show that advertising expenditures are positively related to investor recognition. We scale advertising expenditures by total sales. We follow Green and Jame (2013) in setting missing values to zero and including a dummy variable that equals one when advertising expenditures are missing, and zero otherwise. Loughran and Schultz (2005) document that firms located in urban areas are traded more frequently due to the larger local investor base and easier access to information. Following their approach, we construct dummy variables for firms in urban and rural regions. As in their study, this leaves many stocks unclassified as either urban or rural.

Media coverage is likely to proxy for the amount of information disseminated to a broader audience and for increased investor recognition (e.g., Hillert, Jacobs, and Müller, 2014). We thus construct a dummy variable that is one if a stock was covered by the national media in the previous 12 months and zero otherwise. Turnover also contains an industry component induced by, for instance, style-based investment strategies or differences in industry saliency (e.g., Hong and Stein, 2007; Chen, Hou, and Stulz, 2015). We thus use a set of dummy variables to control for the forty-nine Fama and French (1997) industries.

Our sample period starts in January 1983, when trading volume data for Nasdaq become comprehensively available in CRSP. In addition, some of the control variables (in particular analyst data) are not reliably available in earlier years (see, e.g., Hong, Lim, and Stein, 2000). The sample period ends in December 2011.

2.3 Quantifying the Distribution of Firm Names

We construct two measures, in the following referred to as continuous position and position dummies. Both measures are based on the time-series of (historical and current) company names as provided by CRSP.[5] In each month, we sort all eligible firms on their official company name in alphabetically ascending order. Continuous position simply indicates the relative position of a firm's name in the firm universe under consideration (e.g., NYSE/Amex firms or Nasdaq firms). Consequently, continuous position is uniformly distributed over the interval (0,1]. The position dummies take 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. Position dummy 5 thus 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 one if the firm name is between percentiles 5 and 25. Finally, position dummy 50 (position dummy 75) is one if the name is between percentiles 25 and 50 (50 and 75). Thus, findings are benchmarked against the firms near the end of the alphabetically ordered list (i.e., percentile 75 and above).

The upper half of Figure 1 shows the pooled distribution of firm names by first letter. Findings are benchmarked against the distribution of words in general, which we construct from Google's "Ngram corpus" (see Michel *et al*. (2011) for details). The corpus we rely on is based on the content of hundreds of thousands of digitalized books written in American English and published between 1983 and 2009.

Differences between both word lists are most pronounced at the very beginning of the alphabet; 9.66% of firm names in our baseline sample start with A, which is about 40% higher than the respective relative frequency for words in general. The cumulative distribution functions in the lower half of Figure 1 cross each other between "R" and "Q," which closely corresponds to percentile 75 used as a cut-off point for the position dummies. These insights are in line with anecdotal evidence, which suggests that at least some firms appear to believe in increased visibility due to rank effects.[6]

In order to identify possible determinants of a firm's relative ranking in an alphabetically ordered list, we regress continuous position on all firm characteristics outlined in the previous section. Unless noted otherwise, statistical inference in this article is based on predictive panel regressions. Standard errors are 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. In some specifications, we alternatively rely on Fama and MacBeth (1973)-type regressions, which corresponds to the approach of Chordia, Huh, and Subrahmanyam (2007). To account for serial correlation in the residuals, standard errors are computed using the method of Newey and West (1987)

5 We only consider firms whose name starts with a letter, but additionally including firms whose name starts with a number does not change inferences. In close to 10% of firm months, CRSP reports names whose first letter is followed by a blank space, which we eliminate. Using the unmodified names or excluding these firm months from the analysis does not change inferences.

6 For instance, it is argued that the name "Amiga" was chosen in the 1980s in part to appear before its competitors Apple and Atari (e.g., DeMaria and Wilson, 2002), and that the name "Amazon" was chosen in the 1990s in part to profit from alphabetically ordered web directories (e.g., Byers, 2006). A pattern similar to that in Figure 1 is also often seen in the alphabetical distribution of firm names listed in the Yellow Pages, which leads Einav and Yariv (2006, p. 186) to conclude: "In fact, this sort of influence on attention appears to be heavily exploited in the realm of advertising."

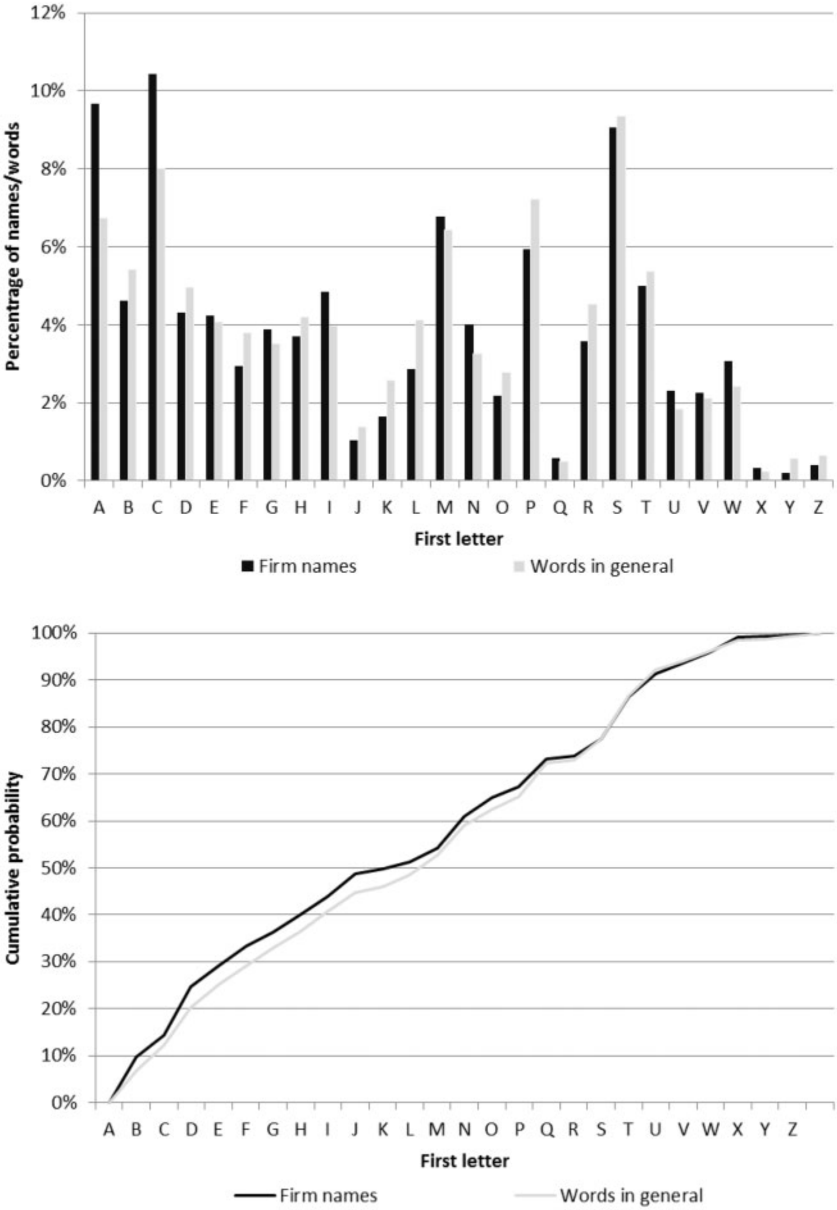

Figure 1. First letter of firm names compared with first letter of words in general.

The upper half of the figure shows the relative frequency of firm names starting with a given letter in comparison to the respective frequencies for words in general. The distribution of firm names is based on all stocks that meet the data requirements for our baseline analysis (see Sections 2.1 and 2.2). The sample period ranges from January 1983 to December 2011. The distribution of words in general is derived from the application "Google Ngram" (see Michel *et al.* (2011) for more details). The corpus is based on the words contained in several hundreds of thousands of digitalized books written in American English and published between 1983 and 2009. The lower half of the figure shows the cumulative distribution functions based on the numbers displayed in the upper half of the figure.

and the automatic lag selection mechanism proposed in Newey and West (1994). Economically, the major difference between both approaches is that the panel regression weights each firm month equally, whereas the Fama and MacBeth (1973) approach weights each month equally. Econometrically, the panel regression approach is likely to be more conservative in the presence of a firm effect (see Petersen, 2009) and thus used as baseline method. Table I shows our findings.

A firm's alphabetic position is essentially unrelated to a large set of firm characteristics. The (adjusted) R^2 from the panel regression is 0.0099 (0.0098), the corresponding numbers for the Fama and MacBeth (1973)-type regression are 0.0295 (0.0019). The only finding that is statistically significant at the 1% level in both approaches is that stocks located at the beginning of an alphabetically ordered list have a higher likelihood of being covered by the media. These firms are not only more visible to investors, but perhaps to journalists as well.

2.4 Baseline Results

Panel A(B) of Table II presents the main findings with regard to name ordering-induced cross-sectional differences in turnover (illiquidity). The central insight is that continuous position indeed matters for the cross-section of both turnover and Amihud (2002) illiquidity. For NYSE/Amex (Nasdaq) stocks, the implied turnover increase for firms at the beginning of an alphabetical listing relative to their counterparts at the end is, all else equal, about 10% (9%). Across exchanges, the turnover difference is 10%, with a t-statistic of 4.40. For the Amihud (2002) illiquidity ratio, the findings are similar. NYSE/Amex (Nasdaq) stocks early in the alphabetical listing are estimated to have about 8% (11%) lower Amihud (2002) illiquidity. Across exchanges, the respective value is 9%, with a t-statistic of 3.04. Our findings with respect to the impact of the control variables are broadly in line with insights from previous work (e.g., Chordia, Huh, and Subrahmanyam, 2007).

Table III reports findings when we alternatively rely on the position dummies. Estimates from the NYSE/Amex/Nasdaq universe indicate that the first 5% of firms in an alphabetical list have, all else equal, about 12% higher turnover and 13% lower Amihud (2002) illiquidity than their counterparts with names in percentiles 75 and above. The effect becomes smaller for firms later in the alphabet. For instance, while firms with names in percentiles 5–25 still have about 8% higher turnover and about 6% lower illiquidity, there is hardly any effect for firms with names in percentiles 50–75. In sum, the findings strongly support our insights from Table II.

3. Evidence from Alternative Settings in Stock Markets

3.1 International Stock Markets

To additionally validate the conjecture that investors facing a number of investment alternatives concentrate more on the alternatives typically offered first, we turn to international stock markets. More specifically, we contrast the pooled trading behavior with respect to alphabetical ordering in the five largest non-Asian international stock markets (Canada, UK, France, Germany, and Australia) with the trading behavior in Japan.

In contrast to the Western countries, the presentation format in Japan is typically determined by numerical ticker symbols, and not by the English alphabet. Stock codes are issued by Japan's national numbering agency and can generally be understood to be assigned based on the order when the stock was listed. Moreover, the ordering contains a strong

Table I. Multivariate relationship between a firm's alphabetic position and other firm characteristics

This table shows coefficients and *t*-statistics (in parentheses) obtained from panel regressions (in Panel A) or from Fama/MacBeth regressions (in Panel B). In both cases, the dependent variable is continuous position, defined as the relative position of a firm's name within the alphabetically sorted firm universe (NYSE/Amex/Nasdaq). The sample period is January 1983 to December 2011. Details on the explanatory variables are provided in Section 2.2 as well as in Table 1 of the Online Appendix. In Panel A, standard errors are double-clustered by firm and month. In Panel B, standard errors are computed using the method of Newey and West (1987) and the automatic lag selection mechanism proposed in Newey and West (1994). In both regressions, statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

	Panel A: Panel regression		Panel B: Fama/MacBeth regression	
Positive return	−0.001	(−0.19)	0.002	(0.54)
Negative return	0.004	(0.54)	0.003	(0.47)
Beta	0.001	(0.74)	0.002	(1.42)
Share price	−0.007	(−1.45)	−0.005	(−1.54)
Age	0.006	(1.14)	0.005*	(1.73)
Market capitalization	0.009**	(2.17)	0.009***	(3.35)
Book-to-market	0.006*	(1.68)	0.008***	(2.92)
Leverage	−0.004	(−0.31)	−0.003	(−0.60)
Earnings volatility	−0.015	(−0.75)	0.005	(0.20)
Earnings surprise	−0.004	(−0.45)	−0.001	(−0.09)
Number of analysts	−0.006	(−1.11)	−0.008**	(−2.05)
Analyst forecast dispersion	0.003	(0.82)	0.002	(0.84)
Advertising	−0.026	(−0.19)	−0.029	(−0.40)
Missing advertising	−0.009	(−1.15)	−0.008**	(−2.06)
Momentum formation return	0.001	(0.67)	0.002	(1.18)
52-week high	−0.001	(−0.55)	−0.003	(−1.19)
Urban firm	0.003	(0.34)	0.002	(1.40)
Rural firm	−0.008	(−0.45)	−0.010**	(−2.29)
Idiosyncratic volatility	−0.015	(−0.40)	−0.001	(−0.02)
Alpha	−0.031	(−0.45)	−0.049	(−0.89)
Return 12 months ago	0.000	(0.21)	0.001	(0.64)
Media coverage	−0.015***	(−3.34)	−0.017***	(−8.79)
S&P 500 membership	−0.032**	(−2.29)	−0.031***	(−6.87)
Dow Jones 30 membership	−0.097*	(−1.96)	−0.098***	(−15.18)
Nasdaq membership	0.004	(0.50)	0.008***	(3.54)
Fama French 49 industry dummies	Yes		Yes	
Year dummies	Yes		No	
(Average) N	956,023		2,747	
(Average) R^2	0.0099		0.0295	
(Average) Adjusted R^2	0.0098		0.0019	

Table II. Stock-level turnover, illiquidity, and alphabetic bias: Predictive panel regressions

This table shows the coefficients and *t*-statistics (in parentheses) obtained from six predictive multivariate panel regressions. The sample period is January 1983 to December 2011. The regressions differ in the dependent variable (Panel A: logarithmized stock-level turnover in month *t*, Panel B: logarithmized stock-level Amihud (2002) illiquidity ratio in month *t*) as well as in the firm universe (NYSE/Amex, Nasdaq, all). In all regressions, the explanatory variable of interest is continuous position (depicted in bold), defined as the relative position of a firm's name within the alphabetically sorted firm universe in the previous month. All control variables are lagged at least by 1 month. If we run regressions across all exchanges, trading volume for NYSE/Amex firms is doubled, and a Nasdaq dummy variable is included. Details on control variables are provided in Section 2.2 and in Table 1 of the Online Appendix. In all regressions, standard errors are double-clustered by firm and month. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

	Panel A: Turnover			Panel B: Amihud illiquidity		
	NYSE/ Amex	Nasdaq	All	NYSE/ Amex	Nasdaq	All
Continuous position	**−0.102*****	**−0.094*****	**−0.101*****	**0.082****	**0.111*****	**0.093*****
	(−3.20)	**(−3.25)**	**(−4.40)**	**(1.98)**	**(2.83)**	**(3.04)**
Positive return	1.010***	0.925***	0.988***	−0.600***	−0.707***	−0.700***
	(16.47)	(15.53)	(16.77)	(−11.28)	(−11.79)	(−12.80)
Negative return	−1.573***	−1.250***	−1.445***	0.445***	0.404***	0.516***
	(−16.68)	(−20.29)	(−21.87)	(4.40)	(4.81)	(6.16)
Beta	0.110***	0.070***	0.092***	−0.082***	−0.073***	−0.088***
	(12.73)	(11.42)	(16.51)	(−9.20)	(−9.73)	(−14.53)
Share price	0.233***	0.024	0.124***	−0.523***	−0.113***	−0.314***
	(10.08)	(1.47)	(9.01)	(−17.01)	(−5.46)	(−17.63)
Age	−0.026**	−0.140***	−0.109***	0.006	0.125***	0.079***
	(−2.13)	(−9.16)	(−10.70)	(0.37)	(5.36)	(5.59)
Market capitalization	0.039**	0.173***	0.117***	−1.059***	−1.330***	−1.208***
	(2.45)	(13.83)	(10.36)	(−54.68)	(−71.34)	(−78.90)
Book-to-market	0.014	−0.080***	−0.043***	0.006	0.142***	0.093***
	(1.01)	(−6.31)	(−4.01)	(0.36)	(8.49)	(7.21)
Leverage	0.459***	0.092**	0.231***	−0.363***	−0.033	−0.121***
	(9.14)	(2.47)	(7.10)	(−6.10)	(−0.68)	(−2.96)
Earnings volatility	−0.343***	−0.136*	−0.197***	0.309***	0.184**	0.208***
	(−3.64)	(−1.83)	(−3.10)	(2.72)	(1.98)	(2.72)
Earnings surprise	0.469***	0.491***	0.486***	−0.223***	−0.364***	−0.336***
	(9.01)	(11.44)	(13.01)	(−3.47)	(−6.98)	(−7.79)
Number of analysts	0.410***	0.519***	0.430***	−0.413***	−0.566***	−0.462***
	(18.07)	(31.19)	(29.01)	(−14.76)	(−23.84)	(−22.89)
Analyst forecast dispersion	0.087***	0.020*	0.053***	−0.067***	0.021	−0.020
	(5.13)	(1.86)	(5.48)	(−3.31)	(1.33)	(−1.49)
Advertising	0.743*	0.019	0.217	0.059	0.500	0.571
	(1.72)	(0.05)	(0.69)	(0.08)	(0.89)	(1.13)
Missing advertising	−0.033	−0.031	−0.034**	0.029	0.036	0.036*
	(−1.41)	(−1.63)	(−2.13)	(0.93)	(1.32)	(1.65)
Momentum formation return	0.064***	0.134***	0.115***	−0.082***	−0.166***	−0.133***
	(3.93)	(9.00)	(9.15)	(−4.88)	(−6.46)	(−7.21)
52-week high	−0.027***	0.032**	−0.020*	−0.063***	−0.236***	−0.123***
	(−2.66)	(2.22)	(−1.81)	(−5.17)	(−12.48)	(−8.99)

(continued)

Table II. Continued

	Panel A: Turnover			Panel B: Amihud illiquidity		
	NYSE/ Amex	Nasdaq	All	NYSE/ Amex	Nasdaq	All
Urban firm	0.022	0.055***	0.039***	−0.020	0.001	−0.014
	(1.21)	(3.05)	(2.83)	(−0.79)	(0.03)	(−0.75)
Rural firm	0.054	−0.089**	−0.026	−0.139***	0.159***	0.021
	(1.44)	(−2.31)	(−0.87)	(−3.19)	(2.93)	(0.53)
Idiosyncratic volatility	3.463***	2.168***	2.514***	−2.219***	−2.076***	−2.249***
	(11.74)	(9.82)	(12.45)	(−6.92)	(−9.01)	(−11.40)
Alpha	−0.186	0.957***	0.520**	0.757*	−1.510***	−0.637**
	(−0.50)	(3.74)	(2.20)	(1.76)	(−4.45)	(−2.25)
Return 12 months ago	0.063**	0.080***	0.076***	−0.047	−0.081***	−0.064**
	(2.56)	(4.23)	(3.79)	(−1.61)	(−3.11)	(−2.54)
Media coverage	0.050***	0.038***	0.034***	−0.083***	−0.079***	−0.068***
	(4.69)	(3.23)	(3.55)	(−6.32)	(−5.05)	(−6.10)
S&P 500 membership	−0.121***	−0.324***	−0.244***	0.037	0.479***	0.279***
	(−4.42)	(−6.29)	(−9.22)	(0.98)	(5.57)	(7.17)
Dow Jones 30 membership	−0.437***	−1.141***	−0.587***	0.363***	1.395***	0.627***
	(−4.84)	(−12.81)	(−5.99)	(3.69)	(12.27)	(5.59)
Nasdaq membership			−0.393***			0.744***
			(−20.88)			(30.30)
Fama French 49 industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	450,661	505,362	956,023	450,599	503,978	954,577
R^2	0.495	0.441	0.449	0.918	0.847	0.894

industry component. We control for these effects by assessing the relative rank of a stock based on the numerical stock code for each industry separately and by including control variables such as age, market capitalization, and the number of analysts.[7] The Japanese version of continuous position turns out to be essentially uncorrelated with continuous position as implied by alphabetically ordered firm names. By construction, it is also uncorrelated with industry classification, which we measure using the Datastream level 4 industry classification.

Control variables are motivated by the model used in the US market, but are restricted to those variables that are broadly available for international stocks. Based on data from Datastream, Worldscope, and I/B/E/S, we control for firm size, firm age, stock price, book-to-market ratio, leverage, beta, 52 week high, idiosyncratic volatility, alpha, momentum formation period return, return 12 months ago, analyst coverage, and analyst forecast dispersion. Construction details are provided in the Online Appendix. All predictive panel regressions also include dummies for years and industries. In contrast to the US market,

7 We use logs of the aforementioned variables in order to be consistent with our analysis for the USA and the other stock markets, and because a regression specification with the logged variables unconditionally explains a higher fraction of trading volume and liquidity. Using raw values does not change inferences.

Table III. Turnover, (il)liquidity, and alphabetic bias: Alternative measure of alphabetic position

This table summarizes the main results of six predictive panel regressions which differ in the dependent variable (Panel A: logarithmized stock-level turnover in month *t*. Panel B: logarithmized stock-level Amihud (2002) illiquidity ratio in month *t*) as well as in the firm universe (NYSE/Amex, Nasdaq, all). Displayed are the coefficients and *t*-statistics (in parentheses) for four dummy variables that indicate the position of a given firm name within the respective firm universe. Position dummy 5 takes on a value of 1 if the company name is among the first 5% in a given month and is zero otherwise. Position dummy 25 is one if the firm name is between percentiles 5 and 25. Finally, position dummy 50 (position dummy 75) is one if the name is between percentiles 25 and 50 (50 and 75). The (untabulated) control variables (including dummy variables for years and industries) correspond to the ones used in Table II. Standard errors are double-clustered by firm and month. The last four rows of each panel provide *p*-values from *F*-tests of coefficient equality. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Turnover as dependent variable

	NYSE/Amex	Nasdaq	All
Position dummy 5 (1)	0.101**	0.108***	0.118***
	(2.34)	(2.82)	(3.74)
Position dummy 25 (2)	0.087***	0.060**	0.076***
	(3.28)	(2.41)	(3.92)
Position dummy 50 (3)	0.059**	0.035	0.050***
	(2.26)	(1.45)	(2.69)
Position dummy 75 (4)	0.044*	−0.008	0.018
	(1.74)	(−0.34)	(1.00)
N	450,661	505,362	956,023
R^2	0.495	0.441	0.449
p-value (1) = (2)/(3)/(4)/0	0.73/0.30/0.17/0.02**	0.21/0.05*/0.00***/0.00***	0.18/0.03**/0.00***/0.00***
p-value (2) = (3)/(4)/0	0.24/0.07*/0.00***	0.29/0.00***/0.00**	0.16/0.00***/0.00***
p-value (3) = (4)/0	0.50/0.02**	0.06*/0.15	0.07*/0.01***
p-value (4) = 0	0.08*	0.73	0.32

Panel B: Amihud illiquidity as dependent variable

	NYSE/Amex	Nasdaq	All
Position dummy 5 (1)	−0.084*	−0.150***	−0.126***
	(−1.73)	(−3.07)	(−3.29)
Position dummy 25 (2)	−0.054	−0.063*	−0.055**
	(−1.49)	(−1.84)	(−2.08)
Position dummy 50 (3)	−0.061*	−0.024	−0.042*
	(−1.83)	(−0.69)	(−1.66)
Position dummy 75 (4)	−0.033	0.041	0.010
	(−1.03)	(1.25)	(0.40)
N	450,599	503,978	954,577
R^2	0.918	0.847	0.894
p-value (1) = (2)/(3)/(4)/0	0.55/0.63/0.28/0.08*	0.08*/0.01***/0.00***/0.00***	0.07*/0.03**/0.00***/0.00***
p-value (2) = (3)/(4)/0	0.84/0.55/0.14	0.25/0.00***/0.07*	0.62/0.01**/0.04**
p-value (3) = (4)/0	0.35/0.07*	0.05**/0.41	0.03**/0.10*
p-value (4) = 0	0.30	0.21	0.70

Table IV. International evidence on ordering effects: Predictive panel regressions

This table explores ordering effects in the cross-section of monthly stock-level turnover (Panel A) or stock-level Amihud illiquidity ratio (Panel B) in pooled Western stock markets (Australia, Canada, France, Germany, UK) and in the Japanese stock market. We gather daily stock market data from Datastream, accounting data from Worldscope, and analyst data from I/B/E/S. The sample period ranges from January 2004 to December 2013. Details about the sample construction are provided in the Online Appendix. Displayed are the main results of predictive panel regressions of the natural logarithm of stock-level turnover or illiquidity on a set of lagged control variables and a lagged measure of the relative position of a stock in a given ordering scheme. The latter is based either on alphabetic ordering of firm names (as of December 2013) or, in the case of Japan, based on the local stock code rank assigned by the Securities Identification Code Committee. More specifically, for each month and each Datastream level 4 industry, we compute the relative position of a firm's stock code. Thus, both the alphabetic ordering measure and the stock code ordering measure are uniformly distributed over the interval (0,1]. We refer to the resulting variables as continuous position. Moreover, position dummies (separately for each ordering scheme) are defined as in Table III. Lagged controls include two variables for the return in the previous month, firm size, firm age, stock price, book-to-market ratio, leverage, beta, fifty-two week high, idiosyncratic volatility, alpha, momentum formation period return, return 12 months ago, analyst coverage, and analyst forecast dispersion. All regressions also include year and industry dummies. In the case of the Western stock markets, we pool the data, add country fixed effects, and also interact these effects with the control variables (except the industry and year dummies). In all regressions, standard errors are double-clustered by firm and month. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Logarithmized monthly turnover as dependent variable

Countries Ordering scheme	Western countries Alphabetic ordering	Japan Alphabetic ordering	Japan Stock code ordering
N	248,850	269,587	269,587
R^2	0.67	0.50	0.50
Panel A1: Impact of continuous position			
Coefficient of continuous position	−0.122***	0.010	−0.196***
	(−3.29)	(0.29)	(−5.40)
Panel A2: Impact of position dummies			
Coefficient of position dummy 5	0.150***	0.043	0.195***
	(3.95)	(1.10)	(3.52)
Coefficient of position dummy 25	0.072**	−0.034	0.123***
	(2.25)	(−1.18)	(3.99)
Coefficient of position dummy 50	0.032	0.006	0.079***
	(1.09)	(0.20)	(2.90)
Coefficient of position dummy 75	0.007	−0.049*	0.045*
	(0.24)	(−1.79)	(1.74)

(continued)

Table IV. Continued

Panel B: Logarithmized monthly Amihud illiquidity ratio as dependent variable			
Countries Ordering scheme	Western countries Alphabetic ordering	Japan Alphabetic ordering	Japan Stock code ordering
N	248,850	269,587	269,587
R^2	0.89	0.87	0.87
Panel B1: Impact of continuous position			
Coefficient of continuous position	0.121***	−0.012	0.141***
	(2.86)	(−0.33)	(3.75)
Panel B2: Impact of position dummies			
Coefficient of position dummy 5	−0.181***	−0.030	−0.171***
	(−3.91)	(−0.73)	(−3.02)
Coefficient of position dummy 25	−0.078**	0.035	−0.080**
	(−2.11)	(1.15)	(−2.50)
Coefficient of position dummy 50	−0.029	−0.005	−0.049*
	(−0.86)	(−0.17)	(−1.73)
Coefficient of position dummy 75	−0.011	0.052*	−0.022
	(−0.31)	(1.75)	(−0.80)

reliable and comprehensive stock-level time-series of historical names do not seem to be available for international stock markets. We thus focus on the most recent 10-year period (January 2004 to December 2013), and base our analysis on the current names. The noise introduced by this approach runs counter to finding significant rank effects. Table IV shows the main findings.

In Table IV, both with respect to turnover (Panel A) and illiquidity (Panel B) and with respect to both continuous position and position dummies, the estimates for the pooled Western countries are similar to those obtained for the US market. Differences between firms early and late in the alphabetical list are in the area of 10% and highly statistically significant.

In contrast, and as hypothesized, we do not find evidence for alphabetic bias in Japan. However, and again as predicted, the implied turnover (liquidity) difference between firms at the beginning and the end of the numerical stock code list is about 20% (15%). As in the US stock market and the Western countries, further tests with position dummies confirm that the firms at the very beginning of a typical stock list profit the most. Collectively, the analysis of international stock markets provides support for the idea of rank effects in stock selection. As inferences are the same irrespective of whether we use continuous position or position dummies, and in order to conserve space and to facilitate presentation, all remaining stock market tests rely on continuous position.

3.2 US Name Changes

Name changes allow us to focus on within-firm variation in alphabetical ranking. Despite their conceptual appeal, name changes are endogenous corporate decisions, and we cannot rule out the possibility of changes in firm characteristics other than names affecting our

Table V. Turnover, (il)liquidity, and alphabetic bias: Name changes

This table summarizes the main results of the alphabetic distribution of firm name changes (in Panel A) as well as of individual firm fixed effect panel regressions (in Panels B and C), which explore the consequences of name changes for stock-level turnover and liquidity. Panel A shows the fraction of name changes that lead to a position earlier in the alphabet. *P*-values (in parentheses) are computed from one-sided binomial probability tests (against a test proportion of 50%). In addition, Panel A shows the average change in continuous position in the context of name changes. *T*-statistics (in parentheses) are computed from *t*-tests with the null hypothesis that the average change is zero. In Panels B and C, we run panel regressions like those in the baseline analysis (Table II). In Panel B (C), the dependent variable is the natural logarithm of monthly stock turnover (the logarithmized monthly Amihud (2002) illiquidity ratio). We require at least twenty-four non-missing monthly observations in the 60-month periods both before and after a ranking change of at least 25%. In addition, we exclude the month of the name change. The (untabulated) control variables correspond to the ones used in Table II. In addition, we include a name change dummy which is one (zero) for the period after (before) the name change. We average continuous position both before and after each event in order to make firm fixed effects capture the impact of material name changes (only). In the same regression, small month to month changes in continuous position are captured by small change, which is one (minus one) in the case of a position earlier (later) in the alphabetical ranking and zero otherwise. Standard errors are double-clustered by firm and month, and year dummies are included in the regression. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Alphabetic distribution of name changes

Strength of name changes	Number of name changes	% earlier in the alphabetic ranking (*p*-value)	Average change in continuous position (*t*-stat.)
Absolute change of at least 25%	750	51.6% (0.20)	−2.72% (−1.40)
Absolute change of at least 50%	326	53.1% (0.15)	−5.32% (−1.40)
Absolute change of at least 75%	91	57.1% (0.10)	−12.53% (−1.42)
Absolute change of at least 90%	14	71.4%* (0.09)	−39.29% (−1.68)

Panel B: Name changes (firm fixed effect regressions) with turnover as dependent variable

Coefficient of continuous position	−0.124*	Coefficient of small change	0.005
t-Stat of continuous position	−1.84	*t*-Stat of small change	0.93

Panel C: Name changes (firm fixed effect regressions) with Amihud illiquidity as dependent variable

Coefficient of continuous position	0.155	Coefficient of small change	0.003
t-Stat of continuous position	1.62	*t*-Stat of small change	0.51

results. However, at least unconditionally, the relative position of a firm's name in an alphabetical listing appears to be reasonably random (Table I). We only consider the 750 name changes that go along with an absolute change in continuous position of at least 0.25. As Panel A of Table V shows, to some extent there seem to be patterns that complement the analysis of the alphabetic distribution of names presented in Figure 1: extreme name changes tend to lead to a position earlier in the alphabetical ranking.

Small within-firm variation in continuous position is often induced by variation in the eligible firm universe, name changes of other firms, or name changes that change continuous position by less than 25%. We do not expect such small changes to strongly affect trading activity. In an attempt to separate these effects from material name changes, we first require at least twenty-four non-missing monthly observations in the 60-month periods both before and after a ranking change of at least 25%. We then average continuous position both before and after each event in order to make firm-fixed effects capture the impact of material name changes (only). Small month to month changes in continuous position are captured by small change, which is one (minus one) in the case of a position earlier (later) in the alphabetical ranking and zero otherwise. We then run panel regressions as before. However, we exclude the month of the name change to control for short-term confounding effects. In addition, we include firm-fixed effects as well as a dummy variable that is one (zero) in the period after (before) the name change.

Our central finding, displayed in Panels B and C of Table V, is that the insights from the between-firm baseline analysis tend to carry over to the within-firm analysis. With turnover (liquidity) as the dependent variable, the coefficient of continuous position is -0.155, with a t-statistic of -1.62 (0.124, with a t-statistic of 1.84). Economically, the effect size is comparable to the cross-sectional findings. As expected, the estimates for small change in the same regression are economically and statistically negligible. Finally, excluding all firm months used in the name change analysis from the cross-sectional tests does not change any of the insights obtained in the baseline analysis.

4. Cross-Sectional Determinants

We start by exploring the hypothesis that rank effects are strongest among otherwise less visible firms. As proxies for a firm's visibility, we consider lagged market capitalization, lagged residual analyst coverage, and lagged industry market share. Following Hong, Lim, and Stein (2000), residual analyst coverage is defined as the residual obtained from monthly cross-sectional regressions of log(1+ number of analysts providing fiscal year one earnings estimates in $t-1$) on log(firm size in $t-1$). We use the forty-nine Fama and French (1997) industries to define industry membership.[8] As industry market share is not based on firm size, but based on relative sales updated once every year at the end of June, each of our three proxies should capture a different aspect of firm visibility. We thus partition our sample along these three dimensions. More specifically, with respect to each proxy and for each month separately, we perform a median split and then run our baseline analysis as in Table II separately for visible and less visible firms. The findings in Panel A of Table VI are in line with our predictions.

In all regressions, both the point estimates for continuous position and the corresponding t-statistics are lower for visible firms than for less visible firms. For instance, small firms (firms with low residual analyst coverage, and firms with low industry market share) positioned early in an alphabetical listing are estimated to have all else equal about 12.5% (13.6%, and 13.5%) higher turnover than corresponding firms found later in an alphabetical listing. The corresponding estimate for large firms (firms with high residual analyst coverage, and firms with high industry market share) is about 6% (7%, and 6%). Findings are similar in the case of liquidity. Firms that have a name within the top of an alphabetical

8 Using the two-digit SIC classification generates similar results.

Table VI. Cross-sectional determinants of the impact of alphabetic bias

This table explores the role of firm visibility (in Panel A) and investor sophistication (in Panels B and C) for the impact of alphabetic bias on turnover and liquidity. We run predictive panel regressions, and we rely on the same controls and sample period (unless noted otherwise) as in the baseline analysis (Table II). In Panels A (firm visibility) and B (investor sophistication), the table shows coefficient estimates of continuous position for different subsamples of stocks based on monthly median splits with respect to the following variables. In Panel A, small firms (large firms) are firms with below (above) median market capitalization in the previous month. To quantify firms with low (high) residual analyst coverage, we rely on the residual from monthly cross-sectional regressions of log(1 + number of analysts providing fiscal year 1 earnings estimates in the previous month) on log(firm size in the previous month). Industry market share is defined as the firm's total lagged sales divided by the total lagged sales of all firms within the market segment under consideration, as defined by the forty-nine Fama and French (1997) industries. Sales are updated once every year at the end of June. To compute aggregate visibility in a given month, we first rank firms with respect to each of the three firm visibility proxies (firms size, analyst coverage, industry market share) in ascending order. Thus, in each month, firms are assigned a value between 0 and 1. We then compute the sum of the three rank measures and perform a monthly median split on the resulting aggregate visibility score. In Panel B, high (low) institutional ownership refers to firms with above (below) median institutional ownership lagged by one quarter. In Panel C, we compare retail turnover based on data from a large discount broker (based on the log ratio of the monthly number of shares traded and the number of shares outstanding) as well as fund turnover (based on the log ratio of the quarterly number of shares traded, as estimated from changes in quarterly holdings, and the number of shares outstanding) with stock-level turnover based on overall market trading activity. Both in the case of retail turnover and fund turnover, missing firm months are considered to have zero turnover, and turnover is winsorized at the 99th percentile. Standard errors are double-clustered by firm and month. Additionally, year dummies are included in the regression. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Median splits based on firm visibility

Subsample	Logarithmized turnover		Logarithmized illiquidity	
Small stocks	−0.125***	(−4.26)	0.134***	(3.51)
Large stocks	−0.060**	(−2.29)	0.046	(1.22)
Difference	0.065*	(1.74)	−0.088*	(−1.76)
Low residual analyst coverage	−0.136***	(−4.20)	0.117***	(2.70)
High residual analyst coverage	−0.071***	(−3.14)	0.071**	(2.35)
Difference	0.065*	(1.90)	−0.046	(−1.02)
Low industry market share	−0.135***	(−4.51)	0.138***	(3.52)
High industry market share	−0.057**	(−2.01)	0.043	(1.06)
Difference	0.078**	(1.98)	−0.095*	(−1.76)
Low aggregate visibility	−0.148***	(−4.79)	0.146***	(3.65)
High aggregate visibility	−0.041	(−1.60)	0.033	(0.92)
Difference	0.107***	(2.85)	−0.113**	(−2.31)
Panel B: Median splits based on investor sophistication				
Low institutional ownership	−0.148***	(−4.63)	0.146***	(3.62)
High institutional ownership	−0.031	(−1.25)	0.022	(0.58)
Difference	0.118***	(3.07)	−0.124**	(−2.41)
Panel C: Alphabetic bias among different groups of market participants				
Retail investors (1991–96)	−0.174*	(−1.75)	(Market: −0.096** (−2.54))	
Fund managers (1983–2011)	−0.067	(−1.22)	(Market: −0.101*** (−4.40))	

listing and are otherwise less visible are about 11.7–13.8% more liquid than corresponding firms at the end of an alphabetically ordered list. For visible firms, the effect is about 4.3–7.1% and in two out of three cases statistically not distinguishable from zero. The meaningful economic differences between the coefficient estimates for visible and less visible firms are statistically significant at the 10% level in all three (two out of three) regressions with turnover (Amihud illiquidity) as the dependent variable.

In order to reduce the noise and condense the information contained in firm size, residual analyst coverage, and industry market share, we also construct an aggregate visibility score. In each month and for each firm, we first compute the relative rank with respect to each of the three variables in ascending order. We then compute the sum of the three rank measures and perform a monthly median split on the resulting score. Panel A of Table VI shows that this procedure is powerful. For less visible firms, the implied turnover (illiquidity) difference between firms early and late in the alphabetical listing is 14.8%, with a *t*-statistic of 4.79 (14.6%, with a *t*-statistic of 3.65). For less visible firms, these estimates are not statistically different from zero. The effect difference is economically meaningful (slightly larger than 10%) and statistically significant (*t*-statistic 2.85 for turnover and 2.31 for illiquidity).

Another plausible conjecture is that our findings may be most pronounced for stocks with a high fraction of unsophisticated investors, who are widely thought of as exhibiting more biases, as facing more cognitive constraints, and as having fewer resources than other market participants. We explore this hypothesis in Panels B and C of Table VI. In Panel B, we again run subsample tests based on a median split. As a natural proxy for investor sophistication, we distinguish between firms with above and below median lagged institutional ownership. To this end, we gather quarterly ownership information for institutional managers with at least 100 million USD in assets under management, as obtained via the Thomson-Reuters Institutional Holdings (13F) Database. The results strongly confirm our predictions. Alphabetic bias leaves discernible traces in stocks with low institutional ownership only. With respect to the impact of continuous position, the difference between both subgroups is similar in size, both economically and statistically, to the difference between firms with below and above median aggregate visibility.

In Panel C of Table VI, we switch from the firm to the investor perspective by changing the dependent variable to retail turnover or fund turnover. More specifically, we compute monthly retail turnover based on data from a large discount broker for the 1991–96 period (see Barber and Odean (2000) and Kumar (2009) for more information). Similarly, we compute quarterly fund turnover based on the Thomson-Reuters Mutual Fund Holdings Database for the 1983–2011 period.[9]

9 Mutual fund holdings are available on a quarterly basis only, which implies that the fund turnover analysis misses within-quarter round-trip stock transactions. Judging from the insights of Table I, there is little reason to suspect that these transactions strongly differ, with respect to alphabetic bias, from other trades. In addition, Puckett and Yan (2011) estimate that within-quarter round-trip transactions account for less than a quarter of mutual funds' trades. In essence, having access to quarterly estimates only may lead to a noisy, but relatively unbiased estimate. Under this assumption, comparing the magnitude of quarterly estimates of alphabetic bias (funds) to monthly estimates (market, retail investors) appears justified as we are interested in relative (but not absolute) cross-sectional turnover differences for a given investor universe.

We then use these turnover measures as dependent variables in our regressions, which otherwise are identical to our baseline approach. With respect to retail turnover, the coefficient on continuous position is only marginally significant but economically meaningful. More precisely, it indicates that firms early in an alphabetical listing have about 17% more retail trading activity. The corresponding estimate for unconditional trading activity over the same time period is only close to 10% for the overall market. In contrast, mutual fund managers' alphabetic bias appears to be about a third smaller than the estimate for the overall market. Under the assumption that mutual fund managers (retail investors) represent relatively (un)skilled investors, these findings again suggest that alphabetic bias is, all else equal, negatively related to investor sophistication.

5. Robustness Checks

Table VII shows the main findings from fourteen robustness tests of our baseline analysis. In order to gain additional insight, we run these tests for all NYSE/Amex/Nasdaq stocks, as well as separately for the subsample of visible and less visible firms, respectively. The classification is based on a median split as determined by the aggregate visibility proxy described in Table VI. In line with our predictions, the impact of alphabetic bias is much stronger (at least twice as large) for less visible firms than for visible firms in each test.

In Panel A, we use the same approach as in Table II, but modify the following aspects.

Additional controls. There may be additional variables that could be correlated with both alphabetic ordering and trading activity or liquidity. We thus add a control for local media coverage, as well as ten controls for corporate events identified by parsing 8-K filings, as explained in detail in Table 1 of the Online Appendix. In addition, we control for company name fluency (Green and Jame, 2013). For data availability reasons, the sample period for this extended model is February 1995 to December 2009. Specification 1 of Table 7 shows that inferences remain unchanged.

Alternative econometric approach. The baseline approach weights each firm month equally. A plausible alternative is to weight each month equally, which we implement by using Fama and MacBeth (1973)-type regressions. Specification 2 of Table VII shows that inferences remain unchanged.

Alternative treatment of values in the right tail of the distribution. Our findings may be driven by outliers. Specification 3, in which we winsorize the dependent variable at the 95% level, verifies that this is not the case. Excluding the affected firm months from the analysis or winsorizing the data at the 99% level yields very similar findings.

Alternative data transformation. The baseline analysis shows that ordering effects explain relative (i.e., multiplicative) differences in trading variables. To test for the impact on absolute (i.e., additive) differences, we use raw (instead of logarithmized) turnover and Amihud illiquidity in specification 4. To reduce the impact of outliers, we again winsorize the data at the 95% level. The qualitative nature of our findings does not change. For instance, firms early in an alphabetical listing are estimated to have 0.8 percentage points higher monthly turnover. In the subsample of less visible firms, the coefficient for continuous position is similar to the coefficient obtained in the full sample regression. However, unconditionally, less visible stocks have only about half the turnover of visible stocks, so that the relative effect size is much stronger for the former.

Alternative frequency. Names are a relatively stable firm characteristic, making a low frequency analysis a plausible alternative to our baseline approach. Where appropriate, we

Table VII. Alphabetic bias: Robustness checks

This table shows the coefficients and *t*-statistics (in parentheses) obtained from predictive regressions. Unless noted otherwise, the econometric approach (panel regressions with standard errors double-clustered by firm and month), the firm universe (NYSE/Amex/Nasdaq), and the sample period (January 1983 to December 2011) are as in the baseline analysis in Table II. Each test is run for all stocks as well as separately for the subsample of visible and less visible firms, as determined by the aggregate visibility proxy described in Table VI. In Panel A, we use the same dependent variables as in Table II. In (1), we add local media coverage, ten dummies for firm news based on form 8-K filings, and company name fluency to the baseline model. The sample period is February 1995 to December 2009. In (2), we run Fama and MacBeth (1973)-type regressions. Standard errors are computed using the method of Newey and West (1987) and the automatic lag selection mechanism proposed in Newey and West (1994). In (3), we winsorize the dependent variable at the 95% level. In (4), we use raw (instead of logarithmized) values of turnover and Amihud illiquidity. In (6), we compute continuous position based on (lagged) ticker symbols instead of based on firm names. In (7) and (8), we compute the alphabetic position relative to each Fama/French forty-nine industry (two-digit SIC code). In Panel B, we use alternative measures of trading activity or liquidity as dependent variable. In (9), (10), and (11) we use the total monthly dollar trading volume, the total monthly number of shares traded, or the monthly number of trades, respectively. Due to limited data availability, (11) can only be implemented using stocks trading on Nasdaq. In (12), we rely on the fraction of days with a zero return in a given month. In (13), we rely on the monthly average of the high-low spread proxy proposed in Corwin and Schultz (2012). In (14), we rely on the monthly average of daily bid-ask spreads computed as in Chung and Zhang (2014). Data are available from 1993 on. Statistical significance at the 10%, 5% and 1% level is indicated by *, **, and ***, respectively.

ID	Description	Dependent variable	Less visible	All stocks	Visible
Panel A: Sensitivity checks of the baseline analysis					
1	More comprehensive model	Turnover	−0.176***	−0.112***	−0.045
			(−4.69)	(−4.33)	(−1.61)
		Amihud	0.206***	0.111***	0.029
			(4.30)	(3.19)	(0.73)2
2	Fama/MacBeth regression	Turnover	−0.128***	−0.085***	−0.039***
			(−7.55)	(−7.61)	(−4.32)
		Amihud	0.131***	0.081***	0.039***
			(6.10)	(7.61)	(2.73)
3	Values winsorized at 95% level	Turnover	−0.146***	−0.098***	−0.039
			(−4.79)	(−4.38)	(−1.57)
		Amihud	0.131***	0.087***	0.034
			(3.47)	(2.97)	(0.96)
4	Raw values winsorized at 95% level	Turnover	−0.008***	−0.008***	−0.004
			(−3.32)	(−2.94)	(−1.09)
		Amihud	0.334***	0.130**	−0.033
			(3.53)	(2.03)	(−1.06)
5	Yearly frequency	Turnover	−0.143***	−0.100***	−0.047*
			(−4.67)	(−4.44)	(−1.91)
		Amihud	0.152***	0.098***	0.038
			(3.58)	(3.23)	(1.10)

(continued)

Table VII. Continued

ID	Description	Dependent variable	Less visible	All stocks	Visible
6	Ticker symbols	Turnover	−0.129***	−0.091***	−0.046*
			(−4.16)	(−3.95)	(−1.77)
		Amihud	0.131***	0.087***	0.035
			(3.26)	(2.83)	(0.95)
7	Within Forty-nine Fama/ French industries	Turnover	−0.141***	−0.094***	−0.036
			(−4.61)	(−4.15)	(−1.42)
		Amihud	0.145***	0.092***	0.031
			(3.64)	(3.02)	(0.87)
8	Within two-digit SIC codes	Turnover	−0.139***	−0.097***	−0.045*
			(−4.59)	(−4.34)	(−1.84)
		Amihud	0.142***	0.091***	0.037
			(3.57)	(3.04)	(1.04)
Panel B: Alternative measures of trading activity and liquidity					
9		Log dollar volume	−0.146***	−0.098***	−0.038
			(−4.69)	(−4.25)	(−1.49)
10		Log share volume	−0.146***	−0.099***	−0.039
			(−4.72)	(−4.27)	(−1.49)
11		Log number of trades (Nasdaq)	−0.126***	−0.108***	−0.044
			(−3.40)	(−3.63)	(−1.32)
12		Fraction of zero return days	0.680**	0.356**	0.187
			(2.51)	(2.00)	(1.20)
13		Log high–low spread proxy	0.036**	0.013	0.007
			(2.21)	(1.04)	(0.60)
14		Log CRSP bid/ ask spread	0.063***	0.034**	−0.005
			(3.14)	(2.15)	(−0.26)

compute yearly averages of all dependent and independent variables (such as turnover or firm size) or we alternatively use end of year values (for variables such as the S&P 500 dummy or industry membership). Estimates are close, both economically and statistically, to the ones obtained in the baseline analysis.

Alternative measures of alphabetical ranking. Stocks are often sorted by ticker symbol instead of by name. Specification 6 shows that findings do not change if we replicate the analysis with ticker symbols. This result is not surprising as in more than 95% of all firm months, the first letter of the company name equals the first letter of the ticker symbol. Unreported tests aimed at exploiting the cases where names and ticker symbols differ indicate that both are equally relevant from a statistical point of view.

In their stock selection process, investors may apply preceding filter rules. For example, they may focus on specific exchanges, as in our baseline analysis. Industrial boundaries are a further natural candidate for such filter rules. We thus compute continuous position relative to each of the forty-nine Fama and French (1997) industries (in specification 7) or relative to each two-digit SIC code (in specification 8). Again, inferences do not change.

Alternative measures of trading activity and liquidity. Conceptually, trading activity and liquidity have many facets. To test for how well our baseline findings can be generalized, we rerun the analysis with the following six dependent variables (specifications 9–14 in

Panel B of Table VII): log dollar volume, log share volume, log number of trades, fraction of zero return days, log high-low spread (Corwin and Schultz, 2012), and log CRSP bid-ask spread (Chung and Zhang, 2014).[10] Again, the effect is mainly concentrated in less visible firms. More precisely, less visible firms at the beginning of an alphabetically ordered list are estimated to have about 14.6% higher dollar trading volume and share volume, 12.6% more trades, 0.68 percentage points fewer zero return days (with an unconditional mean of 20.4%), 3.6% lower high-low spreads, and 6.3% lower CRSP bid-ask spreads than their peers at the end of the list. *T*-statistics for these estimates range from 2.21 to 4.72. Collectively, these results also indicate that alphabetic bias appears is particularly important for the cost-per-volume channel of liquidity. However, and primarily for less visible stocks, alphabetic bias leaves discernable traces in the cost-per-price channel of liquidity as well.

6. Mutual Fund Flows

The mutual fund industry offers another important setting that allows us to study 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 the decision to well-defined and relatively homogeneous market segments, such as open-end US equity mutual funds with a domestic investment focus, as we do in the analysis below. Averaged across all months in our final sample from January 1992 to December 2012, there are close to 1,500 funds whose combined assets under management total 1.69 trillion USD. For the empirical analysis, we rely on the CRSP Mutual Fund Database. We use the third and fourth character of the CRSP style code to assign each fund month to one of nineteen styles in our sample. We use parsing algorithms and manual screens to exclude index funds. We drop very small funds that never have net total assets of over 5 million USD during the sample period. Finally, we exclude observations if any of the control variables (see below) cannot be computed. To avoid multiple counting, we aggregate all share classes of a fund using MFLINKS.

The dependent variable is the net flow for fund i in month t. Following the consensus in the literature, fund flow is defined as the percentage change in total net assets ($\mathrm{TNA}_{i,t-1}$) that is not driven by the fund's return net of fees ($\mathrm{ret}_{i,t}$):

$$\text{Fund flow}_{i,t} = \frac{\mathrm{TNA}_{i,t} - \mathrm{TNA}_{i,t-1}}{\mathrm{TNA}_{i,t-1}} - \mathrm{ret}_{i,t}. \tag{2}$$

In order to mitigate the impact of outliers and to be closer to the normality assumption, we winsorize the distribution at the 2.5% and 97.5% levels in each month separately. Using alternative cut-offs (1%, 5%, 95%, 99%) does not change inferences. The independent variables of interest are again measures that indicate the position of a given fund name in the universe of eligible funds. Consequently, we rely on continuous position as well as on

10 Construction details are reported in Table 1 of the Online Appendix. All measures have been shown to be useful proxies in previous work. For instance, Lesmond, Ogden, and Trzcinka (1999) provide theoretical arguments and empirical evidence that stock-level liquidity can be proxied for by the proportion of zero return days within a month. Fong, Holden, and Trzcinka (2014, p. 4) find that the Corwin and Schultz (2012) measure as well as the Chung and Zhang (2014) measure "are tied-for-best percent-cost proxies for US research". For all proxies, our findings are robust with respect to measuring the dependent variable in absolute units or in log units.

position dummies, defined as in the stock market analysis. We rely on the fund name reported by CRSP.

The selection of control variables is guided by previous work on the determinants of fund flows. The computation of these variables is explained in detail in the notes of Table VIII. 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., Sirri and Tufano, 1998), we take additional style-specific performance measures into account. Specifically, for each year and each of the nineteen investment objectives separately, we determine the relative performance position of each fund. The resulting 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 related to Green and Jame (2013), we compute the length of the fund name. At the fund family level, we compute the total assets under management and the number of offered funds. We also include a dummy variable for the three largest mutual fund 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 all 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 ordering effects induced by alphabetization of fund names. Table VIII shows the main findings.

We run regressions separately for all funds (specifications 1 and 2), as well as for small funds (specifications 3 and 4) and large funds (specifications 5 and 6), determined by a monthly median split. Specification 1 reveals that the impact of continuous position is as expected: funds with names at the beginning of an alphabetical listing generate, ceteris paribus, about 0.16% higher inflows each month. This effect is statistically significant and comparable to moving from the 50th percentile of fund flows to the 54th percentile.

Specifications 3–6 show that the effect is driven by small funds with names at the very beginning of an alphabetical listing. This result 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 Tufano (1998)). The first 5% of smaller than median funds benefit the most by far. As model 4 shows, they achieve more than 0.5% higher inflows per month (or more than 6% higher inflows per year) than small funds toward the end of an alphabetically ordered list. This highly significant effect is equivalent to moving from the 50th percentile of the distribution of flows for below median-sized sample funds to the 60th percentile. Even in this subsample, the average fund still has about 80 million USD under management. Thus, our findings are important from an economic perspective.

Tables 2 and 3 in the Online Appendix show that inferences remain qualitatively unchanged if we use alternative ways to control for style, fund family, and time effects. They also show that the impact of continuous position on net flows into small funds remains stable if we rely on name changes, using the same approach as for stocks in Table V. In the overall picture, insights from the stock market can thus be transferred to the mutual fund market. Being early in an alphabetical listing can lead to economically substantial benefits for mutual fund companies in that funds may grow faster.

Table VIII. Mutual fund flows and alphabetic bias

This table summarizes the main results of predictive panel regressions. The dependent variable is the monthly fund flow (winsorized at the 2.5% and 97.5% level). The sample period ranges from January 1992 to December 2012. Small funds (large funds) are funds with total net assets below or equal to (above) the monthly median. Continuous position and the position dummies are defined as in the stock market analysis. Fund size is the natural logarithm of 1-month lagged total net assets of the fund. Fund age is ln(age in months). Fund expense ratio and fund turnover are lagged by 1 year and logarithmized. Length of fund name refers to the number of letters, after dropping share-class information and incorporation terms. Fund return is the cumulative return (net of fees) over the previous 12 months. Fund risk is the standard deviation of the previous 12 monthly return observations. Fund performance rank indicates the relative performance of the fund in its market segment (one of nineteen styles identified by the CRSP style code) in a given month. Total net assets and number of funds are lagged by 1 month and logarithmized. Top 3 fund families is a dummy variable which takes on a value of 1 (zero) if the fund belongs to one of the three largest families based on 1-month lagged total net assets. Style growth is the current growth rate of the fund's market segment. All regressions contain style-year fixed effects. Standard errors are double-clustered by fund and month. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Fund universe	All funds		Small funds		Large funds	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Continuous position	−0.162**		−0.267***		−0.0012	
	(−2.51)		(−2.95)		(−0.14)	
Position dummy 5		0.325***		0.515***		0.0978
		(3.72)		(4.31)		(1.03)
Position dummy 25		0.0244		0.0934		−0.0925
		(0.44)		(1.20)		(−1.40)
Position dummy 50		0.0625		0.0321		0.0538
		(1.15)		(0.41)		(0.88)
Position dummy 75		−0.121**		−0.110		−0.153**
		(−2.26)		(−1.55)		(−2.41)
Fund size	−0.0465***	−0.0465***	−0.184***	−0.188***	0.00593	0.0073
	(−2.72)	(−2.73)	(−6.10)	(−6.26)	(0.26)	(0.32)
Fund age	−0.824***	−0.825***	−0.975***	−0.975***	−0.689***	−0.692***
	(−21.86)	(−21.90)	(−19.29)	(−19.28)	(−16.48)	(−16.60)
Fund expense ratio	0.0391	0.0409	0.0271	0.0300	0.0884	0.0959
	(0.64)	(0.68)	(0.36)	(0.40)	(1.10)	(1.19)
Fund turnover	−0.0428*	−0.0413*	−0.0125	−0.0129	−0.0814***	−0.0771***
	(−1.76)	(−1.70)	(−0.41)	(−0.42)	(−2.92)	(−2.77)
Number of share classes	−0.231*	−0.205	−0.181	−0.151	−0.417***	−0.415***
	(−1.83)	(−1.62)	(−1.08)	(−0.90)	(−3.11)	(−3.08)
Length of fund name	−0.0045***	−0.0046***	−0.0075***	−0.0080***	−0.0012	−0.0014
	(−2.92)	(−2.99)	(−3.50)	(−3.70)	(−0.68)	(−0.78)
Fund return	3.977***	3.973***	4.420***	4.414***	3.643***	3.637***
	(10.75)	(10.75)	(10.37)	(10.37)	(10.27)	(10.26)
Fund risk	0.441	0.471	1.995	2.045	−0.320	−0.320
	(0.22)	(0.24)	(0.87)	(0.89)	(−0.15)	(−0.15)
Fund performance rank	0.266	0.268	0.181	0.184	0.459**	0.472**
	(1.54)	(1.55)	(0.82)	(0.84)	(2.38)	(2.46)
(Fund performance rank)^2	0.826***	0.824***	1.074***	1.070***	0.418**	0.404**
	(4.96)	(4.95)	(4.88)	(4.87)	(2.34)	(2.26)

(continued)

Table VIII. Continued

Fund universe	All funds		Small funds		Large funds	
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Fund family: Total net assets	0.173***	0.175***	0.197***	0.200***	0.119***	0.122***
	(7.77)	(7.86)	(6.49)	(6.61)	(3.51)	(3.59)
Fund family: Number of funds	−0.279***	−0.284***	−0.316***	−0.318***	−0.237***	−0.247***
	(−7.25)	(−7.34)	(−5.24)	(−5.27)	(−5.24)	(−5.40)
Top 3 fund families	0.421***	0.409***	0.706***	0.733***	0.387***	0.334***
	(5.27)	(5.01)	(3.75)	(3.87)	(4.49)	(3.73)
Style growth	91.38***	91.39***	91.46***	91.49***	90.82***	90.84***
	(24.06)	(24.07)	(21.46)	(21.48)	(19.79)	(19.80)
N	374,566	374,566	187,359	187,359	187,207	187,207
R^2	0.133	0.133	0.125	0.125	0.164	0.165

7. Conclusion

Sorting names alphabetically is an omnipresent convention. For many settings in academia, economics, and politics, this practice has been shown to yield an advantage to those positioned early in an alphabetical listing. We are the first to analyze implications of this type of alphabetic bias in financial markets. We find that a better alphabetic ranking goes along with higher stock-level trading activity and liquidity. We also find that alphabetic bias is negatively related to firm visibility and investor sophistication. The cross-section of stock-level trading in international markets as well as of mutual fund flows lends further support to the idea that alphabetical ordering has economically important consequences for financial decision making.

It is difficult to provide a convincing rational mechanism for these findings. In a game theory setting, one might argue that investors aim at trading assets that others are trading, because they consider liquidity to be attractive. Alphabetical order could then serve as a focal point (e.g., Schelling, 1960). This line of reasoning could indeed explain why the effects are strongest for assets that are otherwise less visible: focal points are most needed for them. However, this rational explanation stands in contrast to the effects' being weakest for those stocks that are mostly traded by sophisticated investors.

Our findings suggest that the position of a firm's name in an alphabetically ordered list may be used as an instrument for trading activity in future research. Furthermore, our results have economic implications for several actors in financial markets. For instance, they allow investors to identify stocks that, all else equal, are less costly to trade. Our results may be of relevance to firms in search of comparatively cheap methods to increase liquidity. With regard to the presentation and marketing of their product universe, the results might be of interest to index providers, brokers, banks, and mutual fund companies.

Supplementary Material

Supplementary data are available at *Review of Finance* online.

References

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

Ang, J., Chua, A., and Jiang, D. (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 Dyl, E. A. (1997) Market structure and reported trading volume: NASDAQ versus the NYSE, *Journal of Financial Research* **20**, 291–304.

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

Barber, B. M. and Odean, T. (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.

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

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

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

Benartzi, S. and Thaler, R. (2007) Heuristics and biases in retirement savings behavior, *Journal of Economic Perspectives* **21**, 81–104.

Bond, P., Edmans, A., and Goldstein, I. (2012) The real effects of financial markets, *Annual Review of Financial Economics* **4**, 339–360.

Brandt, M., Brav, A., Graham, J. R., and Kumar, A. (2010) Idiosyncratic volatility puzzle: Time trend or speculative episodes?, *Review of Financial Studies* **23**, 863–899.

Brogaard, J., Engelberg, J., and Parsons, C. A. (2014) Networks and productivity: Causal evidence from editor rotations, *Journal of Financial Economics* **111**, 251–270.

Byers, A. (2006) *Jeff Bezos: The Founder of Amazon.com (Internet Career Biographies)*, Rosen Publishing Group.

Campbell, J. Y., Grossman, S. Y., and Wang, J. (1993) Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* **108**, 905–939.

Carney, D. R. and Banaji, M. R. (2012) First is best, *PLoS One* **7**, 1–5.

Chen, H., Noronha, G., and Singal, V. (2004) The price response to S&P 500 index additions and deletions: Evidence of asymmetry and a new explanation, *Journal of Finance* **59**, 1901–1929.

Chen, J., Hou, K., and Stulz, R. M. (2015) Are firms in "boring" industries worth less?, NBER Working Paper 20880.

Choi, D., Getmansky, M., and Tookes, H. (2009) Convertible bond arbitrage, liquidity externalities, and stock prices, *Journal of Financial Economics* **91**, 227–251.

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

Chung, K. H. and Zhang, H. (2014) A simple approximation of intraday spreads using daily data, *Journal of Financial Markets* **17**, 94–120.

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

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

Cooper, M. J., Khorana, A., Osobova, I., Patel, A., and Rau, P. R. (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.

Corwin, S. A. and Schultz, P. (2012) A simple way to estimate bid-ask spreads from daily high and low prices, *Journal of Finance* **67**, 719–759.

Coval, J. and Stafford, E. (2007) Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* **86**, 479–512.

DeMaria, R. and Wilson, J. L. (2002) *High Score! The Illustrated History of Electronic Games*. Osborne/McGraw-Hill.

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

Edmans, A., Goldstein, I., and Jiang, W. (2012) The real effects of financial markets: The impact of prices on takeovers, *Journal of Finance* **67**, 933–971.

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

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

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

Fedenia, M. and Hirschey, M. (2009) The Chipotle paradox, *Journal of Applied Finance* **19**, 144–164.

Feenberg, D. R., Ganguli, I., Gaule, P., and Gruber, J. (2015) It's good to be first: Order bias in reading and citing NBER working papers. Unpublished working paper, National Bureau of Economic Research, CERGE-EI, University of Massachusetts, MIT.

Fong, K. Y., Holden, C. W., and Trzcinka, C. A. (2014) What are the best liquidity proxies for global research? Unpublished working paper, University of New South Wales, Indiana University.

French, K. R. (2008) The cost of active investing, *Journal of Finance* **63**, 1537–1573.

Goetzmann, W. N. and Kumar, A. (2008) Equity portfolio diversification, *Review of Finance* **12**, 433–463.

Goyenko, R., Holden, C. W., and Trzcinka, C. A. (2009) Do liquidity measures measure liquidity?, *Journal of Financial Economics* **92**, 153–181.

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

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

Grinblatt, M. and Keloharju, M. (2001b) What makes investors trade?, *Journal of Finance* **56**, 589–616.

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

Hartzmark, S. M. (2015) The worst, the best, ignoring all the rest: The rank effect and trading behavior, *Review of Financial Studies* **28**, 1024–1059.

Heath, C. and Tversky, A. (1991) Preference and belief: Ambiguity and competence in choice under uncertainty, *Journal of Risk and Uncertainty* **4**, 5–28.

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

Hillert, A., Jacobs, H., and Müller, S. (2014) Media makes momentum, *Review of Financial Studies* **27**, 3467–3501.

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

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

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

Hong, H. G. and Stein, J. C. (1999) A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* **54**, 2143–2184.

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

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

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

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

Keloharju, M., Knpfer, S., and Linnainmaa, J. (2012) Do investors buy what they know? Product market choices and investment decisions, *Review of Financial Studies* **25**, 2921–2958.

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

Lesmond, D. A., Ogden, J. P., and Trzcinka, C. A. (1999) A new estimate of transaction costs, *Review of Financial Studies* **12**, 1113–1141.

Llorente, G., Michaely, R., Saar, G., and Wang, J. (2002) Dynamic volume-return relation of individual stocks, *Review of Financial Studies* **15**, 1005–1047.

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

Loughran, T. and Schultz, P. (2005) Liquidity: Urban versus rural firms, *Journal of Financial Economics* **78**, 341–374.

Meer, J. and Rosen, H. S. (2011) The ABCs of charitable solicitation, *Journal of Public Economics* **95**, 363–371.

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.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Picket, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., and Aiden, E. L. (2011) Quantitative analysis of culture using millions of digitized books, *Science* **331**, 176–182.

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

Newey, W. K. and West, K. D. (1994) Automatic lag selection in covariance matrix estimation, *Review of Economic Studies* **61**, 631–653.

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

Puckett, A. and Yan, X. (2011) The interim trading skills of institutional investors, *Journal of Finance* **66**, 601–633.

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

Richardson, M. L. (2008) Alphabetic bias in the selection of reviewers for the American Journal of Roentgenology, *American Journal of Roentgenology* **191**, 213–216.

Schelling, T. C. (1960) *The Strategy of Conflict*. Harvard University Press, Cambridge.

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

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

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