A COGNITIVE BIAS OF TRADERS IN THE STOCK MARKET

YI WANG

PhD

The Hong Kong Polytechnic University

2020

The Hong Kong Polytechnic University

School of Accounting and Finance

A Cognitive Bias of Traders in the Stock Market

Yi Wang

A dissertation submitted in partial fulfilment

of the requirements for

the degree of Doctor of Philosophy

March 2020

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____________________ (Signed)

_____Yi Wang_________ (Name of student)

A Cognitive Bias of Traders in the Stock Market

ABSTRACT

In this paper, I investigate whether investors in the stock market use round numbers as cognitive reference points during the trading process. Using transaction-level data provided by Abel Noser, I confirm that even for professional institutional investors, their transaction prices tend to cluster at round-number prices. And the degree of the clustering is consistent with the accessibility of the numbers, i.e. whole dollars, half dollars, dimes and nickels, which I refer to as the 'round-number' bias.

Then, within the Abel Noser sample, there is a considerable heterogeneity of the 'round-number' bias: those trades submitted by institutions in larger size groups or those that have specialized trading departments exhibit lesser degree of the bias, comparing with trades from smaller institutions or do not have designated traders inside. Also, building on some findings in prior studies using the same data set, I find that at broker level, those trades executed by 'discount' brokers (who mainly focus on executing trades on behalf of their clients and charge lower commissions) are less clustered at round numbers, comparing with those executed by traditional brokerage houses. These cross-sectional differences indicate that the degree of the bias is related to the efforts devoted by investment companies and brokers to the trade execution process.

Finally, to further study the plausible causes of the 'round-number' bias during the trading stage, I move on to the TAQ data set, which contains most of intraday transactions for securities listed in major stock exchanges in the U.S. Using different weighting methods, I find that on an average trading day, the degree of the 'round-number' bias in the overall market is much higher comparing with that in the Abel Noser sample, suggesting that retail investors are severely affected by such price preference when submitting their orders. And due to the over clustering around 'round number' prices, those 'round-number' trades incur higher trading costs (measured by the commonly used 'effective spread'), ranging from an annual amount of 200 to 900 million of dollars during our sample period (2001-2014). Yet, we do observe a significant declining trend of the 'round-number' bias across time, which can be attributed to the speedy and broad adoption of the algorithm-based trading practice.

Keywords: Institutional Investor; Execution Trader; Cognitive Bias; Algorithm Trading; Behavior Finance; Market Microstructure

Acknowledgments

To begin with, I would like to express my appreciation to my chief supervisor, Dr. Gang HU, for his enthusiastic guidance along my journey at PolyU. Through his words and actions, I have learnt that one can eventually achieve his goal as long as he sets his mind on it and never gives up. Meanwhile, I am very much grateful to Dr. Jing ZHAO for her enduring support and warm encouragement for this project.

Our AF (School of Accounting and Finance) indeed feels like a warm and caring family to me, owing to the substantial and selfless support from Prof. Agnes CHENG, our 'Head'. Frankly, I feel so lucky and fortunate to join the PhD programme here at AF under her leadership. She has devoted all her energy into nourishing and helping the development of AF family and we all benefit tremendously from her generous support, in terms of on-campus conferences, programming and topic-based workshops and social interaction events. I sincerely wish her a good health and happiness going forward.

As Confucius said, "amongst three companions, there must be someone whom you can learn from". Thus, I would like to offer my sincere thanks to my caring and knowledgeable PhD fellows at AF, specially my office-mates in MN043, including Hou Fangfang, Chen Xiaoqi, Yang Zhonghuang and others, for their unconditional encouragements and enjoyable company over the years. Last but not the least, the biggest thank shall be owed to my parents, the 'real' Doctor Wang and Doctor Yan, for their firm supports and caring along my academic expedition. The strong and unwavering self-belief that they have endowed me with will always be my power station to overcome any challenge in the future.

Table of Contents

Chapter 1. Introduction

Economists have long recognized that both individuals and corporate managers have bounded rationality (Simon, 1955; Baker and Wurgler, 2013). A growing literature on human heuristics and cognitive limitations, originating primarily from the field of psychology, has shown that people often use simple cognitive shortcuts when processing information, leading to systematic biases in decision making (Hirshleifer, 2001, 2015). As a commonly applied pricing strategy, companies that sell either consumer or durable goods, such as cameras and refrigerators, often take advantage of consumers' psychological traits and price their products at 9-digit ending prices (e.g. Thomas and Morwitz, 2005; Coulter and Coulter, 2007). And this naïve price preference is even observed in the Swedish housing market, where apartments with asking prices just below round millions receive more bids during the auction, although they are more likely to be overpriced (up to 5%) comparing with similar flats (Repetto and Solis, 2017). In the context of the stock market, a number of studies have documented that stock prices tend to cluster at certain price levels (e.g. $1/2, $1/4) (Harris, 1991; Christie and Stultz, 1994; Ikenberry and Weston, 2007). Yet they provide a variety of possible explanations for this pattern and offer no discussion about the potential impact of such price-clustering phenomenon.

In this paper, I focus on the economic impact of cognitive reference points on a group of market participants that play a key role in the price formation process: the institutional investors. As discussed in Hu (2009), in general, within the same buy-side institution like the hedge fund, a group of profession traders work closely with their portfolio manager on refining their trading strategies and reducing unnecessary transaction costs when executing their managers' orders. Prior literatures mostly focus on the behavioral biases of portfolio managers, and shows that these biases affect their stock-picking and market-timing performance.[1] However, little work has been done to investigate the potential impact of behavioral biases on the trading process of those professional investors in the stock market. A recent article, written by a former professional trader, points out even in the current investment landscape where most

[1] For example, Coval and Shumway (2005) document trading patterns that are consistent with loss aversion among traders in the Chicago Board of Trade. Frazzini (2006) finds evidence of the disposition effect among U.S. mutual fund managers, whereas Barber et al. (2007), using international data, show similar evidences in a broader set of investors. Using a similar dataset, Goetzmann, Kim, Kumar, and Wang (2015) show that the weather-induced mood affects institutional trading and stock returns.

strategies are implemented partially or fully by computerized algorithms, human execution traders still play a vital role[2]. Since the major duty for a professional trader is to get the best possible execution price, a human trader who is affected by some cognitive biases (e.g. the preference for certain numbers with cultural context, see Bhattacharya et al. (2018)) might end up with additional trading costs, which will ultimately affect the performance of the institution as a whole. Thus, we intend to fill this gap and investigate the cognitive biases of both the institutions and the traders they hire[3]. In particular, we conjecture that the cognitive limitation of institutional investors (traders) can lead to the over reliance on round numbers (such as integer or half-dollar prices) as their cognitive reference points, which in turn impacting their trading decisions and trading costs.

Reliance on round numbers as cognitive reference points may lead to two types of trading biases of institutional traders. Firstly, as human beings, institutional traders may also exhibit a preference of round-number prices, where the hierarchy of the roundness follows as: whole dollars, half-dollars, dimes, nickels (which are refered as 'round-number' prices throughout the paper)[4]. Kuo, Lin, Zhao (2015) show that individual investors in Taiwan futures market tend to submit limit orders at round-number prices and they suffer from such 'round-number' bias due to longer order execution time, lower execution rate and innate inferior trading skills. If institutional traders, despite the perception that they should be more sophisticated, are affected by the same 'round-number' bias, we may observe that their trade prices tend to cluster at those round numbers as well.

The second number-related trading bias can be referred to as 'left-digit' bias. The idea is that traders may focus on the leftmost digit of a number while partially ignoring other digits. When investors anchor on a round number, a change in the leftmost digit of a price dramatically affects the perception of the magnitude. For

[2] See the post by Rober Carver (COMMENT: What does a hedge fund execution trader do all day?) at https://news.efinancialcareers.com/hk-en/3001269/what-is-an-execution-trader

[3] Nevertheless, as pointed out by Robert Carver, smaller funds or those newly 'start-up' funds may also rely on human traders (in some cases the portfolio manager itself) given the 'diseconomies of scale' when the marginal benefit of forming an automated trading desk is well below the marginal cost. And in some large institutions, the management team choose multiple routes to execute the orders so that they can evaluate the performance of different trading desk and have a back-up plan when the market is in turmoil.

[4] However, the specific hierarchy of price (number) preference might vary across different countries with different cultural context or historical backgrounds. For instance, using the transaction data from Germany, Fritz (2014) presents evidences of prices clustering at multiples of 20 Euro cents, instead of 25 cents. A series of studies done by Jason Mitchell show that due to Chinese (and some East Asian) investors' preference for the lucky number 8 and unlucky (evil) number 4, the closing prices are more likely to end at prices with the last digit being 8, instead of being 4 (Brown, Chua and Mitchell, 2002; Brown and Mitchell, 2008; Mitchell, 2001).

example, when assessing the increase from $6.99 to $7.00, a trader would anchor on the leftmost digit, which changes from 6 to 7, and believe that it is almost a $1 increase. This pattern has been documented by Bhattacharya, Holden, and Jacobsen (2012) using the Trade and Quote (TAQ) data set, in which they find that there is abnormally high (low) buy-sell imbalance when stock prices are one cent below (above) a 'round-number' price (e.g. '.10', '.20')[5]. Also they find that this type of preference creates temporal buying (selling) pressure for stocks that are priced 1-penny below (above) a 'round-number' price, which cause significant price reversals 24 hours after the current trade.

In fact, these two cognitive biases stem from the same cognitive heuristic, which is relying on round numbers as cognitive short-cuts (and they are not mutually exclusive). A typical trader may exhibit both biases at the same time. We test this conjecture using the Abel Noser institutional trading data (hereafter referred as AN data), which contains detailed intraday transaction records of a subsample of institutional investors in the U.S. Firstly, we find the evidence of the over-representation of 'round-number' prices in the execution prices of the AN sample. Since fund managers are mostly responsible for portfolio decisions, when it comes to order execution, the abnormal clustering of the transaction prices at 'round-number' prices is most likely due to the cognitive biases of the execution traders. To the best of our knowledge, our paper is among the first to uncover the cognitive bias of institutional investors and the traders they hire.

Moreover, we document some important cross-sectional variations of such behavioral bias. In particular, we find that the 'round-number' bias we discovered in the AN sample is much more severe among trades submitted by small-size institutions (using the total in-sample trading volume) and those institutions without designated traders. And trades executed by 'full-service' brokers (i.e. traditional investment banks like *J.P. Morgan*) tend to exhibit larger degree of the 'round-number' bias than those executed at 'discount' brokers (those mainly focus on executing the trade on behalf of their clients, e.g. *Interactive Brokers*). Moreover, we find that this type of behavioral bias is most prominent in the overall market (based on the daily transaction data from the Trade and Quote data set, i.e. the 'TAQ' in abbreviation). These findings suggest

[5] Using the Abel Noser trade sample, we replicate the empirical design of Bhattacharya et al. (2012) and document the existence of such 'left-digit' bias in our sample as well. See the contents in the Appendix section.

that investor's trading sophistication is an important driver for the abnormal price clustering at 'round-numbers'.

We also investigate the impact of these cognitive biases on the trading outcome. Building on the recent development in the empirical market microstructure literature, we estimate the difference in the 'effective spread' measure between trades that are executed at 'round-number' prices and those at other 'non-round-number' prices, using the TAQ sample. Indeed, we find that trades executed at those 'round-number' prices (e.g. '.00', '.50') cost a bit higher in terms of the spread measure of the transaction cost, which can translate to around 785 million of additional trading costs annually within our sample period (2001-2014).

Understanding the role of cognitive biases in institutional traders' trading decisions is important for several reasons. Firstly, as Bai, Philippon, and Savov (2016, p. 627) reports: "Institutional investors have come to dominate financial markets, their stake in the average firm rising from 20% in 1980 to 60% in 2014." Hence, as individual investors delegate more wealth to institutional investors, it is increasingly important to understand what institutional traders do and how they perform. Prior studies have offered a battery of evidences confirming the existence of stock-picking and market-timing skill of institutional investors, especially among active traders like the mutual fund managers with aggressive investment styles or hedge fund managers[6]. For instance, focusing on the quarterly holding of the active mutual funds in the U.S., Chen, Jegadeesh and Wermers (2000) reports that growth-oriented funds exhibit higher level of stock-picking comparing with those income-oriented funds. And follow-up studies have documents that those stocks recently bought by active mutual funds tend to experience positive abnormal returns in the following quarterly earnings announcements (Baker, Litove, Watcher and Wurgler, 2010) and this stock picking skill is more significant during expansion periods (Kacperczyk, Veldkamp and van Nieuwerburgh, 2014).

Second, even if individual investors trade on their own, their investment outcomes are likely to be affected by the aggregate trading behavior of institutional traders, who are also prone to certain types of behavioral biases. As institutional traders are the major market participants in the equity market (Xu, 2015), their trades

[6] Another strand of literature that casts doubt on the 'abnormal' (or 'above-benchmark') returns of the average open-end equity mutual fund in the U.S. include Carhart (1997), Fama and French (2008) etc. For the relevance of this study, we abstain from this line of debating.

are likely to affect the market liquidity and price formation process, which will influence the welfare of all market participants. Recently, a growing strand of literature document the 'irrationality' of institutional investors as a whole, who are commonly regarded as 'sophisticated' or 'informed' traders from both the public and academic points of view. In particular, based on the commonly used Thomson Reuters 13-F data set, Devault, Sias and Starks (2019) discover that the aggregate portfolio allocation decisions of those large institutions are strongly correlated with the market-wide sentiment metrics: they increase their demand for more risky stocks (measured by historical return volatility) when the sentiment starts to rise (based on Baker-Wurgler's index). Another study by Edelen, Ince and Kadlec (2016), which also utilizes the 13-F holding data, presents a series of evidences showing that institutions as a group tend to bet on the 'wrong' side of the previously discovered cross-sectional stock anomalies. And this 'irrationality' is not driven by the flow-induced demand shocks from the retail investors, as previously studied in Frazzini and Lamont (2008). Hence, our study complements this line of research by focusing on a small yet very important facet of those large players in the stock market: their order submission and execution process.

Lastly, the finding in our paper also lends support to the surge in the recent development in algorithmic trading (AT). As discussed in Chordia, Roll, and Subrahmanyam (2011), the drastic increase in the daily trading volume and share turnover rate in the U.S. stock market is partially due to the increasing use of AT by hedge funds and other institutions in recent years. The popularity of those trading programs signals the need for neutral and bias-free trading decisions. Indeed, following prior literature (Hendershott, Jones and Menkveld, 2011) which uses the 2003 'auto-quote' technical improvement event in the NYSE as an event that facilitates the AT, we present some evidences showing that by allowing for more AT activities, the overall degree of the 'round-number' bias in the market is alleviated.

The remaining sections of this study are laid out as follows. First, in Section 2, we conduct a brief overview of the regulatory changes regarding quoting rules in major stock exchanges in the United States. Next, we present our 4 testable hypotheses on the existence, degree and time-series patterns of the 'round-number' bias among our sample of institutional investors or in the overall stock market. Our main results are presented in Section 3 and we summarize our main findings and discuss for potential future explorations in Section 4.

Chapter 2. Institutional Background and Hypotheses Development

2.1. Institutional Background and Literature on Price Clustering

Starting from April, 2001, all U.S. major stock exchanges (namely the NYSE, AMEX and NASDAQ) [7] alter from fractional pricing scheme (both in terms of quoting and trading) to the decimal pricing. Before that, the minimum pricing variation (MPV or commonly referred as 'tick size') of the U.S. equity exchanges has been $1/8 (and $1/16 or $1/32 for low-price stocks) ever since the inception. Closely following this major policy change, numerous studies have shown that both bid-ask spread and the quoted depth are reduced significantly (e.g. Bessembinder, 2003; Chakrabarty and Chung, 2004). Furthermore, in 2005, the SEC adopted Rule 612 of Regulation NMS, imposing the MPV of all listed stocks priced no less than $1.00 to be $0.01, i.e. one penny, and prohibiting "market participants from displaying, ranking, or accepting quotations, orders, or indications of interest in any NMS stock priced in an increment smaller than $0.01 if the quotation, order, or indication of interest is priced equal to or greater than $1.00 per share" (See 17 CFR 242.612; SEC Release No. 34-51808). And for stocks priced below $1.00, the MPV shall equal $0.0001 (so-called 'sub-penny'). Hence, throughout the paper, we first round the transaction price to the penny level and then calculate the 'submission ratio' [8] at a particular 'round-number' price '.XY' (where X and Y are integers ranging from 0 to 9) as the ratio between the number of trades or shares/dollars traded at that particular price and the total number of trades or shares/dollars traded for a sample institution/trader/broker/common stock within a trading horizon (day/month/year). Throughout the paper, we refer to this ratio as the 'round ratio' at a particular 'round-

[7] To be precise, the actual implementation date for all stocks listed in the NYSE/AMEX for decimal pricing is Jan. 29, 2001 and the phase-in date for NASDAQ stocks is Apr. 2, 2001. In order to keep a long time-series sample of Abel Noser trading record, here we include NYSE/AMEX stocks' trading data after Feb. 2001 and NASDAQ stocks after Apr. 2001, as applied in Bhattacharya, Holden and Jacobsen (2012).

[8] Precisely speaking, the correct term shall be the frequency/proportion of trades that are executed at a particular type of rounded penny price being '.XY' (where X and Y are integers ranging from 0 to 9), given that we do not have the dataset that contains the original 'limit orders' that are submitted by portfolio managers/traders, as in Kuo et al. (2015). Nevertheless, since the 'trades' consist a subset of the submitted orders, as long as the proportion of ineffective orders is relatively small or the probability of successful execution is not determined by the distance between the limit order price and the closest 'round-number' price, we believe the 'round-number' bias (if any) discovered based on the executed transaction data shall indicate the cognitive bias of the PM/trader who submitted the order.

number' price '.XY'. And we drop stocks with the month-end closing price below \$1.00 to avoid the 'sub-penny' trading after the Regulation NMS[9].

The phenomenon that both trade and quote prices tend to cluster at certain numbers has long been discovered and investigated in academic literature. To begin with, when analyzing the fixing price of the London gold market, Ball et al. (1985) concluded that the degree of price resolution (similar to the price clustering beyond the quoting rule) is a function of the amount of information, price level and the variability of the asset. Borrowing from the 'costly information acquisition' concept from Grossman and Stiglitz (1980), they argue that due to the inelastic supply of information, the price will be set at coarser grid when the market is volatile. Yet, as pointed out by the author, the London gold market was quite liquid and transparent, with a limited number of participants who might form tacit cooperation (see later in Christie and Schultz, 1994). Later on, focusing on prices of the U.S. common stocks, Harris (1991) re-affirmed the findings in Niederhoffer and Osborne (1966) that prices tend to cluster at even-eighths. In addition, he found that this clustering is more pervasive among stocks listed in the NASDAQ. The author offered a rationale for such a finding, arguing that in a dealer-driven market, market makers use a coarser set of minimum price increments to facilitate negotiation and secure their market making profits.

Apart from the abovementioned 'rational' theories on the price clustering phenomenon, recently, several studies have offered a behavioral explanation, which stems from the innate psychological traits of investors. One prominent study done by Ikenberry and Weston (2007), has found that in their post-decimalization sample (the last 6 months of 2002), the degree of trade prices clustering at '.X0' and '.X5' increased significantly comparing with their pre-decimalization sample (1996). Moreover, they conducted a multivariate analysis by including a set of stock characters as right-hand-side variables and found that a large proportion of clustering at '.X0' and '.X5' cannot be explained by those candidate variables, such as price level and return volatility. Therefore, the authors proposed an additional hypothesis:

[9] Yet, as pointed out by Buti, Rindi, Wen and Werner (2013): "*Rule 612 prohibits market participants from quoting prices in sub-penny, but in the belief that sub-penny trading would not be as detrimental as sub-penny quoting, it expressly allows broker-dealers to provide price improvement to a customer order that results in a sub-penny execution*". And the development of dark pooling markets has made the Rule 612 ineffective at protecting displayed limit orders.

there are some deep-rooted psychological preferences for certain numbers in the decimal system.

In the psychology and marketing literature, numerous studies have found that humans rely on round numbers (e.g. multiples of 5, 10 or 25) to save their cognitive energy and simplify the communication process. For instance, based on a nation-wide sample of newspaper advertisements, Schindler and Kirby (1997) report that there is an overrepresentation of prices ended with 0, 5 and 9 and they relate this finding to the concept of ‘cognitive accessibility’ proposed by Tversky and Kahneman (1979). And under the ‘mental accounting’ theory (Thaler, 1985), a 9-ending price is framed as an integer price plus a small-sized gain. Based on those psychological traits, Bhattacharya et al. (2012) conducted a comprehensive study using a randomly selected sample of the TAQ data from 2001 to 2006. Specifically, they argue that because of the round number heuristics, investors are more likely to buy (sell) if the price of a stock drops (rises) one penny below a round-number threshold (which they term as the ‘left-digit’ effect). And given the prevalent evidence on the clustering of the limit order at round number (e.g. Chung and Chiang, 2006), investors may strategically ‘undercut’ their limit orders to get ahead of the order book. Overall, their results confirm their hypothesis that investors do use round numbers as reference points: there is excessive buying (selling) by liquidity demanders when prices are one penny below integers, half-dollars, quarter, dimes and nickels[10]. Using the limit order submission and execution data from 2004 to 2008, Kuo et al. (2015) document that both individual and institutional investors who exhibit higher degree of ‘round-number’ bias in the previous year tend to have lower mark-to-market returns (ranging from day 0 to day 5). Another related study done by Liu (2011) reports that the intraday spot exchange trades between NTD (New Taiwanese Dollar) and USD tend to cluster at 0-ending and 5-ending prices. And unlike evidences from the stock market, the foreign exchange (FX) markets are dominated by large financial institutions like banks and brokerage houses, who submit large orders and watch the market closely. Other studies documenting the clustering patterns in the FX market

[10] Nevertheless, the specific pattern of the price clustering may be influenced by the local culture, currency, etc. For instance, when analyzing the high-frequency trading data in Germany, Fritz (2014) reports that the sample transaction price tend to cluster at multiples of 20 cents, instead of 25 cents (a quarter), which fits with the values of Euro coins in circulation. Other international studies outside the U.S. also find the clustering pattern is governed by local institutional and cultural factors (e.g., Brown et al., 2002; Brown and Mitchell, 2008; Bhattacharya et al., 2017).

include Goodhart and Curcio (1991), Sopranzetti and Datar (2002), Mitchell and Izan (2006) and etc.

2.2. Hypotheses Development

As summarized in the previous section, so far, there are three main hypotheses regarding the existence of the price clustering in the financial market: namely the 'valuation uncertainty' hypothesis (Ball et al., 1985), the 'cost of negotiation' argument (Harris, 1991) and the behavioral explanation (e.g. Aitken et al., 1996). It is worth of clarifying that those 3 types of hypotheses/explanations are not mutually exclusive. For example, when studying the cross-sectional determinants of the abnormal frequency of trades clustered at '.X0' and '.X5' prices, Ikenberry and Weston (2007) include a set of explanatory variables that are rooted from both the 'valuation uncertainty and the 'negotiation cost' hypothesis. Though they find that the estimated coefficients are in general consistent with the theory's prediction, e.g. the degree of clustering increases with price level, return volatility, average bid-ask spread and average trade size, the economic magnitude of the clustering left unexplained remains very significant. Hence, they conclude that most of the clustering in the stock market is caused by the collective psychological trait of investors who rely on prominent numbers (e.g. 5- or 0-ending numbers) to simplify their trading decisions.

Given in our context where our main research question relates to the existence of institutional investors' cognitive bias, particularly reflected in their order submission and execution process, we propose our first set of hypotheses as below.

H1: Institutional investors exhibit the 'round-number' bias in their daily trading activities (a); yet, the degree of such bias shall be lower comparing with the 'average' investor in the market (b).

To further study the potential impact of institutions' trading experience or the sophistication of their trading department on their 'round-number' bias, we explore the vast heterogeneity of the institutions in the Abel Noser data. Specifically, a study done by Goldstein et al. (2008) has shown that the distribution of the per share commission in this data set has remained quite stable: over 50% of the trades were charged with 5 or 6 cents per share. And they attribute the robust institution-broker relation to the 'soft-dollar' hypothesis made by Blume (1993) that 'full-service' brokers such as J.P. Morgan that offer bundled services to their clients (e.g. timely analyst reports and IPO allocations). Another study done by Anand et al. (2012) who also utilize the Abel Noser data set finds that institutional investors with larger trading volume tend to have lower percentage trading cost. Hence, we explore the heterogeneity of the Abel Noser data to further investigate the effect of institutional investors' trading experience or sophistication on their 'round-number' bias. Specifically, our second set of hypotheses is laid out below:

H2: The degree of this cognitive bias decreases with the sophistication of the trading desk and the trading experience of institutional investors, specifically:

Institutions in larger sizes shall exhibit lesser degree of the 'round-number' bias comparing with their counterparties in smaller sizes (a); institutions whose fund managers are also responsible for the trading shall exhibit higher degree of the 'round-number' bias comparing with those institutions with designated traders (b); institutions who use 'discount' brokers shall exhibit lesser extent of the 'round-number' bias comparing with those who hire the full-service brokerage house (c).

Next, we move on to study the current landscape of the trading venue, namely the algorithmic trading (AT). As noted in an early paper by Jain (2005), the majority of exchanges in each country have adopted computerized order-matching engines that facilitate AT. For instance, using one-month trading data of the Deutscher Aktien Index (DAX) 30 stocks, Hendershott and Riordan (2013) report that AT accounts for around 50% of the marketable limit orders in Jan., 2008. Numerous recent market microstructure studies have shown quite consistent evidences supporting the upside of the AT or the High-Frequency-Trading (HFT)[11]. To name a few, using a stratified sample of trading data with identities of HFT firms from the NASDAQ, Brogaard, Hendershott and Riordan (2014) conducted a comprehensive study showing that trades submitted by HFT firms help with the price discovery process by trading in the direction of permanent price changes. Early in 2003, the NYSE implemented a computer-based 'auto-quote' system for those inside-quote orders. In particular, Hendershott et al. (2011) conclude that the staggered phase-in of such order-matching engine significantly facilitates the use of the AT since the number of messages per minute increase dramatically afterwards. Thus, we exploit such an exogenous event to study the impact of AT on the 'round-number' bias in the market:

H3: The degree of the 'round-number' bias in the market shall be reduced across

[11] As noted by Brogaard et al. (2014), HFT belongs to a category of AT. For more detailed description, please refer to the survey paper by Biais and Woolley (2011).

time when more algorithm-based trading strategies are adopted in the market.

Existing studies on tick size issues have provided unanimous evidences showing that the quoted or effective spread has declined significantly when the tick size reduced (e.g. Ahn, Cao and Choe, 1996; Bessembinder, 1999; Zhao and Chung, 2006). But, on the meantime, numerous studies also find that the market depth (measured by the depth of the limit order book) declined as well. A comprehensive study done by Goldstein and Kavajecz (2000) that analyzes the impact of the NYSE's switching from 1/8 to 1/16 tick size in 1997 find that the average the quoted depth has declined by 48% (similar findings can be found in Huang and Stoll (1997) and Chung, Van Ness and Van Ness (2002)). Hence the overall effect on investors trading cost is ambiguous, depending on the size and types (market vs. limit order) of the trade (e.g. Goldstein and Kavajecz (2000)). For instance, using the sample of institutional trades provided by the Plexus Group, Jones and Lipson (2001) report that the implementation cost of their sample institutions has increased dramatically, especially for those liquidity demanders such as momentum traders.

Given that in our study, we focus on the clustering effect of trades executed around 'round-number' prices, real-time trading cost measures (e.g. the effective spread or price impact measures) are affected both by the size of quoted gap as well as the depth of the limit order book. Intuitively speaking, if investors in general are affected by such cognitive bias and submit their orders at 0- or 5-ending prices, then the depth of the order book would increase more comparing with other digit-ending prices and would take more time to be executed. Indeed, a recent study done by Alexander and Peterson (2007) has found that trades in the NYSE and NASDAQ tend to cluster at multiples of 500, 1000 and 5000 shares and those clustered trades tend to have larger trading costs, measured by effective spread and the price impact. Hence, we present our last hypothesis to be tested as below:

H4: Due to the over clustering of the trades at the 'round-number' price, the transaction cost would be higher for those 'round-number' trades, comparing with the 'non-round-number' trades.

Chapter 3. Data Source Description and Sample Selection

In order to answer the aforementioned research questions, in this study, we rely on two main sources of data sets: the NYSE Trade and Quote (TAQ) data set (2001-2014) and the daily transaction record of its institutional clients provided by Abel Noser (2001-2011). Though the latter data set is anonymized by Abel Noser to protect the trading strategy of its clientele, the data provider has assigned unique IDs (*clientcode*) to each institution who has subscribed its trading cost analysis services. Yet, starting from Oct. 2011, the *clientcode* has been removed from this data set which makes us unable to track the trading performance of each sample institution.

Following previous empirical literature on market microstructure (e.g., Hasbrouck, 2009; Bhattacharya, Holden and Jacobsen, 2012), we construct our TAQ sample in the following steps: first, we match the trading 'symbol' in the TAQ master files (stored by month in the TAQ monthly product) with the 'permno' code in the CRSP monthly return files based on the first 8-digit of historical CUSIPs provided by the CRSP MSE file; second, we restrict our sample to those common stocks listed in one of three major exchanges in the U.S. (NYSE/AMEX/NASDAQ) and did not change their listing locations within a calendar year; then, we drop those stocks whose month-ending prices were below $1 or above $1000, similar to the filter applied by Hendershott, Jones and Menkveld (2011). Due to the fact that the TAQ stops providing the monthly product after 2014 and then switch to the daily product of which the master file is not always available, we end our TAQ sample on Dec. 31, 2014.

[Insert Table 1 here]

Following those previous microstructure studies (e.g. Hasbrouck, 2009; Goyenko et al., 2009), we apply a series of filters to address for the estimation error caused by withdrawn quotes and those trades/quotes in abnormal conditions. First, we only include trades and quotes during normal trading hours: from 9:30 am to 4:05 pm in Eastern Time [12]. Second, we drop those intraday trades labelled with irregular conditions, as in Holden and Jacobsen (2014). Then, following the common practice in microstructure studies, we drop NBBO quotes in non-normal conditions, or with the bid-ask spread larger than $5, or the quoted spread is zero/negative, or the ratio

[12] According to the Foote 2 in Chordia et al. (2001), intraday trades are allowed to report to the Consolidated Tape up to 5 minutes after the core trading session ends.

between the quoted spread and the price is larger than 40%, or the ratio between the effective spread (price minus the midpoint of the NBBO quote) and the bid-ask spread is larger than 4, as applied in Chordia, Roll and Subrahmanyam (2001). After applying the previous filtering criteria, we end up with a sample of common stocks with valid trading data in the TAQ database, ranging from over 5,600 stocks in 2001 to 2,678 stocks in 2014 (reported in Table 1). Yet, the intensity of the trading activity has increased gradually in recent years, from around 3 million trades per day in 2001 to 30 million of trades per day in our last sample year. One the other hand, the average trade size has declined from over 16,000 dollars per trade in 2003 to around 8,000 dollars in 2014, which echoes the observation made by Chordia, Roll and Subrahmanyam (2011) and other recent microstructure studies that the algorithmic trading has become much more rampant in recent decades.

Our daily institutional trading data come from Abel Noser. It is a large propriety data set that contains transaction-level data on buy and sell orders from institutional investment managers (e.g., Fidelity Investments) and pension plan sponsors (e.g., California Public Employees' Retirement System (CalPERS)). Trading data for investment managers are delivered directly via their Order Delivery System. According to Hu et al. (2018), this data set has received increasing popularity among academic researchers aimed at exploring various types of research questions in accounting, investment and other fields. As argued in Anand et al. (2012), because those institutions subscribe to Abel Noser for its execution cost analysis, we believe the *price* variable in this dataset is the actual execution price[13].

Before we move to present the summary statistics of our sample institutions and the stocks they traded, it is necessary to first describe a typical trading process occurred at a buy-side institution and different responsibilities of investment managers and their dedicated traders (if any). As in the Figure 1 of Hu et al. (2018), a portfolio manager (PM) sends an order pertaining to buying or selling a particular stock within a certain range of quantity and trading horizon (typically within one

[13] Existing literature that utilize Abel Noser dataset in their main analyses have shown that comparing with an average institution in the Thomson Reuters 13F dataset, the Abel Noser sample institutions are relatively larger, both in terms of average number of stocks held and the dollar trading volume per quarter (Anand et al., 2011). Since by nature those institutions care about their execution quality so that they pay for the service provided by Abel Noser, we believe using such a 'sub-sample' of the universe of institutional investors is appropriate: if the result confirms our hypothesis (H1) that those 'large players' exhibit 'round-number' bias when executing their trades, then it is more plausible that for other players of smaller sizes such kind of heuristic bias also exists.

trading day) to a trader inside the same institution. Then, the trader is supposed to send the order to external trading venues (e.g. ECN) or brokers and execute at the best price possible[14]. Therefore, the heuristic bias documented by our study (if any) is more of a manifesto of the cognitive limitation of professional traders (and broker-dealers), rather than the management team of a buy-side institution.

In Panel B of Table 1, we report the summary statistics of the daily transaction record from the Abel Noser sample[15]. Here, we report the daily average of the number of transactions/unique institutions/brokers/'traders'[16] within each calendar year. One average, there are around 400 unique institutions sending their trading records to Abel Noser with a sample trading day. Yet, their trading size has declined gradually from over 300,000 dollars in the beginning year of our Abel Noser sample (hereafter refered as 'AN sample) to less than 10,000 dollars per day in 2011, similar to the finding in Cready, Kumas and Subas (2014) that the sharp decline in the average trade size after 2005 is observed only among those institutions with large trading volume during a calendar year. Another pattern worthy of noticing is that the number of unique clients in our AN sample starts to decline after 2006, which is a manifesto of its declining clientele due to the fast changing landscape of the trading-cost-analytics (TCA) industry.

[14] It is worth of pointing out that not all buy-side institutions separate the trading duty from their fund managers. For instance, in some online job posts, candidates are required to have certain past experience in derivative trading and the trading process per se (https://www.optrust.com/CareerOpportunities/Portfolio-Manager-Multi-Strategy-Investments-February-2020.asp).

[15] As suggested by Hu et al. (2018), here we only focus on transactions that are submitted by institutions with their client type codes equal to 1 or 2, i.e. pension plan sponsors or investment managers, respectively. And we use the 'brokercode' as the identifier of brokers that are in charge of executing the trades of Abel Noser's clients.

[16] Here, we only include those 'valid' trader codes that are not purely numeric or missing, as matched by the '*clienttrdcode*' via the 'TraderXref' table provided by Abel Noser.

Chapter 4. Empirical Analyses

4.1. Evidences of the 'Round-number' Bias among Institutional Investors

Following the approach applied by Puckett and Yan (2011) when estimating the coverage of Abel Noser trading volume as a percentage of the CRSP volume, we calculate the percentage of the AN sample trades that are executed at certain types of prices in the following steps. First, for each sample transaction, we round the execution price to the nearest penny level and remove the digits before the decimal point[17]. Then, for each sample stock/institution, we count the number of trades executed at a certain 'rounded penny' price (e.g. '.50') within a sample trading day and scale by the total number of trades for the same stock/institution within the same day. And we term this ratio as 'round ratio (N)' throughout the paper. Meanwhile, we also compute this ratio by summing the number of shares or dollars traded at a certain 'rounded penny' price first and then scale by the share or dollar volume traded for the same stock (or by the same institution) within the same day. The latter two are named as 'round ratio (S)' and 'round ratio (D)', respectively. Finally, we can compute the equal-weighted (EW) mean of this 'round ratio' across all sample stocks/institutions within the same day or calculate the value-weighted (VW) average of those ratios based on the total number of trades (N) / shares (S) / dollars (D) traded for the sample stock or by the sample institution within the same trading day. As mentioned in our Chapter 2 related to the quoting rule during our sample period, we believe such rounding method can capture the 'rounding' heuristic in investors' minds (if any) in the post-decimalization era.

In Figure 1, we plot the time-series average of the VW 'round ratio (N&S&D)' of trades executed at different 'rounded penny' prices across all sample institutions within our AN sample period (Apr. 1, 2001 to Sep. 30, 2011). The figure demonstrates that on an average trading day, the percentage of trades executed at certain 'rounded penny' price (e.g. '.15') relative to the total number of trades/shares or dollars traded across all AN sample institutions within the same day.

[Insert Figure 1 here]

---

[17] As mentioned in Section 3, the *price* variable in this data set is the actual transaction price, not the original limit order price. Yet, as pointed out by Boehmer, Jones and Zhang (2017), institutional investors are heavy users of crossing networks and 'dark pools' where they do not receive much price improvement as those offered by brokerage houses to their retail clients. Hence, we believe that any evidence of the overly clustering of trades executed at 'round-number' prices based on this data set could lend support to our 1st hypothesis that institutional investors are influenced by such cognitive bias.

From Figure 1, apparently one could find that the most favored prices for a sample Abel Noser institution are prices ended with '.00', i.e. the integer, whether in terms of the number of trades (Panel A)/dollars (Panel B)/shares (Panel C) traded. The second most preferred prices are 'half-dollar' prices: those ended with '.50'. For an average institution from AN sample, the daily percentage of trades executed at '.50' is over 1.5%. And for prices that are multiples of 'dime', e.g. '.10', '.20', their average 'round ratio' are larger than that ended with '.X5' (where X is an integer ranging from 1 to 9, e.g. '.15', '25' and etc.). Nevertheless, even for the latter type of 'round-number' prices, their average 'round ratio' is slightly larger than 1%, the unconditional frequency under the discrete uniform distribution. To formally test whether the empirical distribution is different from the discrete uniform distribution, following Ikenberry and Weston (2007), we conduct the Pearson's Chi-squared test supplemented with the Kolmogorov–Smirnov test in Table 2. The latter one is more suitable for the continuous distribution but can be applied for two-sample comparison. The results strongly reject the null hypothesis that the expected frequency of trades executed at a 'rounded penny' price equals 1% on an average trading day in our AN sample. Thus, both the graphical and univariate statistical analysis have shown that among our sample institutions, there exists a clear pattern of 'round-number' bias based on their transaction prices: more than 2% of the trades are executed at the 'rounded dollar', i.e. '.00', followed by the 'half-dollar' price, i.e. '.50', then the 'dimes', e.g. '.20', '.60' and finally those 'nickels', e.g. '.25', '.35' and etc. And this hierarchy across different types of 'round number' prices is consistent with the cognitive accessibility, as discussed by Kuo, Lin and Zhao (2015)[18].

[Insert Table 2 (Panel ABC) here]

In addition, to better estimate the degree of such 'round-number' bias in our AN sample, we conduct the following regression analysis for each trading day by grouping the 'rounded penny' price according to the roundness of the price, as in the following model:

$$round\,ratio|^{"}.XY^{"}_{t} = \beta_0 + \beta_1 \cdot D_{00} + \beta_2 \cdot D_{50} + \beta_3 \cdot D_{X0} + \beta_4 \cdot D_{X5} + \varepsilon_{XY,t} \qquad (1)$$

where $round\,ratio|^{"}.XY^{"}_{t}$ equals the 'round ratio' of the sample TAQ trades executed at the rounded penny price '.XY' during the trading day t (where X and Y

[18] Since the sum of these 'submission ratios' has to equal 1, those 'round number' prices other than the abovementioned four types of prices will end up with the average 'submission ratio' below 1%, which will be formally estimated in the multivariate regression setting in Section 4.2.

are integers from 0 to 9); D_{X0} groups dime prices such as '.10', '.20' together and D_{X5} represents rounded penny prices like '.15', '.25' and etc. (i.e. the nickels). Throughout the paper, we refer those 4 types of 'rounded penny' prices as 'round-number' prices. Here, the regression coefficients ($\beta_1, \beta_2, \beta_3, \beta_4$) measure the degree of trades clustering at certain type of 'round number' price comparing with the average 'round ratio' of trades executed at prices other than '.00', '.50', '.X0' and '.X5'. For each trading day, we run the OLS regression using Equation (1) and compute the time-series average of the parameter estimated above. In order to control for possible autocorrelations in the error term, the standard deviations of the parameter are calculated using Newey-West (1987) with 5 lags.

[Insert Figure 2 here]

Our next step is to formally test our Hypothesis 1(b): though professional money managers exhibit this behavioral bias when executing their orders, the magnitude of such 'round-number' bias is lower than that of the retail investor. Since the TAQ data set includes almost all the trades and quotes records for all securities listed in the U.S., trades submitted by retail and institutional investors are both recorded in the dataset[19]. Given the mixed nature of this dataset, our Hypothesis 1(b) predicts that the degree of 'round-number' bias inferred from the TAQ trades shall be more significant that that based on AN sample of institutional trades. Therefore, we run a series of daily regression during each sample trading day using the following specification and report the time-series average of the coefficient (reported in the Column 3 of Table 4).

$$\begin{aligned} 'round\ ratio'|'.XY'_t = {} & \beta_1 \cdot D_{00} + \beta_2 \cdot D_{50} + \beta_3 \cdot D_{X0} + \beta_4 \cdot D_{X5} + \beta_5 \cdot D_{Others} \\ & + \beta_6 \cdot D_{00} \times Indicator + \beta_7 \cdot D_{50} \times Indicator + \beta_8 \cdot D_{X0} \times Indicator \\ & + \beta_9 \cdot D_{X5} \times Indicator + \beta_{10} \cdot D_{Others} \times Indicator + \varepsilon_{XY,t} \end{aligned} \tag{2}$$

where the 'Indicator' equals 1 if the 'round ratio' of the rounded penny price '.XY' is calculated based on our AN sample. Here, we simply combine the daily price-level 'round ratio' data sets based on the AN sample and the TAQ sample into one data set and run the daily regression (from Apr. 2001 to Sep. 2011). The estimated coefficients of those interaction terms (β_6 to β_9) are consistent with our Hypothesis 1(b) that this 'round-number' bias due to human's heuristics is relatively less severe for institutional investors, comparing with the overall market trading activities. Thus,

[19] A recent study done by O'Hara, Yao and Ye (2014) documents that those 'odd-lot' trades, with shares less than 100, were not recorded by the TAQ database. Later the SEC ordered exchanges to report the odd-lots to the consolidated tape, starting from Dec.9, 2013.

both the graphical and regression analysis offer consistent evidences supporting our Hypothesis 1(b) that the degree of 'round-number' bias is indeed more prominent among a mixed sample of trades submitted both by retail and institutional investors (i.e. the TAQ sample).

[Insert Table 4 here]

In Table 4, we report the regression results and the F-tests that compare the difference among those coefficients. For those non-'round-number' prices, the percentage of trades executed is slightly lower than 1%, which indicates that the 'round ratio' for other prices should be larger than the unconditional mean of the discrete uniform distribution of 100 variables. Consistent with the figure shown in Figure 1, there is a strict hierarchy of the 'round ratio' among different types of round-number prices, according to the roundness of price. And the F-tests show that indeed the 'round ratio' is highest for trades executed at whole number prices, '.00', followed by that of half-dollars, dimes and nickels, respectively. Hence, both the graphical and regression analyses confirm our 1st hypothesis that investors do rely on 'round-number' prices as cognitive shortcuts when making buying/selling decisions.

4.2. Difference of the 'Round-number' Bias across Different Types of Institutional Investors

In order to further investigate that whether for our AN sample institutions, those who put more emphasis to their trading department, either by forming a specialized trading desk or routing more orders to those 'niche' brokers that focus more on reducing transaction costs (e.g. ECNs or 'dark pools'), exhibit lesser degree of the 'round-number' bias in their executed trades (H2), we look into our AN sample in details.

First, as described in the Figure 3 of Hu (2009), those designated professional traders may exhibit lesser degree of 'round-number' bias due to their long-term experience in trade execution. Here, we utilize the identifier of 'traders' provided in a separate reference file by Abel Noser. Specifically, we separate our sample into two subgroups based on whether a trade can be matched with a 'valid' 'trader' code from the 'TraderXref' table via the 'clienttdrcode'. Here, 'valid' 'trader' codes refer to those contain alphabetical characters and are not purely numeric (e.g. 'Jack', '). Then, we apply the same steps in Sec. 4.1 when computing the 'round ratio' at each 'rounded penny' price in our AN sample and plot the time-series average in Figure 3A and 3B, for trades with 'trader' codes and those without, respectively. From those two figures, one can find that the average ‘round ratio (N)’ of the trades executed at different hierarchy of ‘roundness’ prices, e.g. ‘.00’, ‘.50’ and multiples of one dime (e.g. ‘.20’, ‘.30’), among trades executed by identifiable ‘traders’ from AN sample is slightly lower comparing with the rest of the sample transactions. And the formal statistics in the Panel A of Table 3 strongly reject the equal distribution hypothesis, which lends supports to our Hypothesis 2(a) that the ‘round-number’ bias is decreasing with the sophistication and experience of institutional investors in terms of executing trades.

[Insert Figure 3A & 3B here]

Second, some prior empirical studies have shown that in recent decades (post-2003), traditional ‘full-service’ brokers (those who provide ‘bundled’ services including trade execution, analyst research, underwriting business and etc.) began to face increasing challenges from ‘new entrants’: crossing networks, ECNs and ‘discount’ brokers such as *TD Ameritrade* who only provide execution services with less than half of the commission (even zero commission fees) charged by the traditional brokerage houses. For instance, also based on the Abel Noser data set, Goldstein, Irvine, Kandel and Wiener (2009) document that the volume of trades executed at the discount commission (less than 3 cents per share) takes up over 40% of all the sample volume in 2003. Another study by Anand et al. (2012) shows that the trading cost of ‘execution-only’ brokers is relatively lower comparing with ‘full-service’ brokerages. Since the former specialize in trade execution process, it is possible that orders executed by this type of brokerage (or trading venues) contain less ‘round-number’ bias, as argued in our Hypothesis 2(b). Following Goldstein et al. (2009), we separate our AN sample trades into two groups: those that are more likely to be executed by ‘discount’ brokers (whose in-sample dollar-weighted per share commission is less than 3 cents, identified by the ‘brokercode’) and the rest (more likely to be executed by ‘full-service’ brokerages). And we plot the T-S average of the ‘round ratio (N)’ for different ‘rounded penny’ prices within each subsample in Figure 4A and 4B, respectively. The figures indicate that the degree of the ‘round-number’ bias is more severe among trades executed by the ‘full-service’ brokerage. And the formal statistical test in Panel B of Table 3 also confirms our conjecture (H2b) that ‘execution-only’ brokers do exhibit lesser degree of ‘round-number’ bias comparing with their ‘full-service’ counterparties.

[Insert Figure 4A & 4B here]

Lastly, institutional investors may also differ in their trading experience simply due to the difference in the size of their assets-under-management (AUM). Generally speaking, the total trading cost (including the explicit and implicit components) of large orders tend to exceed those of small orders (Keim and Madhavan, 1996). Yet, it is not necessarily true for large institutional money managers. For instance, a recent study done by Busse, Chordia, Jiang and Tang (2019) reports that after controlling for the difference in investment styles, the percentage execution cost of larger active equity mutual funds is lower comparing with their smaller counterparties. Moreover, numerous studies in the investment literature have confirmed the positive flow-performance relation both at fund and fund-family level (e.g. Jiang and Yuksel, 2017; Nanda, Wang and Zheng, 2004), which leads to the growth of the outperforming funds (Berk and van Binsbergen, 2015). Hence, we conjecture that larger institutions tend to exhibit lesser degree of the 'round-number' bias (H2(c)). To test this prediction, we first aggregate the dollar trading volume of our AN sample institutions from Apr. 2001 to Sep. 2011. Then those 'clients' whose total dollar volume fall into the top (bottom) 33% in-sample are labelled as 'Big' ('Small') institutions. Next, we compute the 'round ratio (N&S&D)' separately within those two subsets of trades that are submitted either by those 'Big' or 'Small' institutions. The histogram in Figure 5 and the formal statistics in Panel C of Table 3 partially confirm our H2(c), as those coefficients of the interaction terms in Column 3 of Table 5 are not significant at common statistical level, except for the coefficient of '.00', which shows that on an average trading day in our sample, the percentage of trades from those large sample institutions executed at '.00' prices is lower by 0.26% comparing with that of these trades from small institutions. Because the daily 'round ratio' is computed under the value-weighting approach, it is mostly affected by those sample institutions with large AUMs or active trading strategies.

[Insert Figure 5A & 5B here]

[Insert Table 3 (Panel ABC) here]

Besides, following Kuo, Lin and Zhao (2015), we perform the following cross-sectional regressions using the Equation 2, in order to estimate the marginal difference of the 'round ratio (N)' across different types of subsamples of the Abel Noser data, by running a series of daily regressions and compute the time-series average of the coefficient estimated. '$Indicator$' is an interaction term that equals 1 if the 'round ratio (N)' is calculated based on the trades of those with matched 'valid' 'trader' codes (Column 1 in Table 5), or if those trades are executed by brokers with per share commission less than 3 cents (Column 2 in Table 5), or if the trades are submitted by those 'Big' institutions in our sample (Colum 3 in Table 5). For each trading day, we run a cross-sectional regression of Equation 2 for 200 different 'rounded penny' prices and then we report the time-series average of the coefficient and the Newey-West adjusted *t*-statistics (by 5-lags) in Table 5.

[Insert Table 5 here]

In column 1 of the table above, we perform the similar OLS procedure as Eq. (2) and formally test the difference of the 'round-number' bias between trades executed by those sample AN institutions with 'valid' trader codes and those without, as defined when plotting Figure 3. The estimated coefficients of those interaction terms (β_6 to β_9) are all negative and significant (the *p*-value is well below 1%). For column 2 and 3, the estimated coefficients of the interaction terms are mostly negative and significant, lending support to our H2(b) and H2(c) that trades executed by those 'discount' brokers, e.g. ECNs, execution-only brokers, and trades executed by larger institutions exhibit lesser degree of 'round-number' bias, comparing with their counterparties. Though only the coefficient of the integer price (D_{00}) in column 3 is significantly negative, the magnitude of the coefficient is quite large: for an average trading day, the percentage of trades executed at an 'integer' price among those institutions with larger trading volume is lower by 0.26% comparing with those from smaller players. Because we use the total in-sample dollar trading volume as a rough proxy of the institution's AUM, the results in column 3 shall be interpreted with caution. Overall, both the univariate analysis in Table 3 and the detailed daily regression estimates in Table 5 are in line with our Hypothesis 2 conjecturing that this type of behavioral bias is less severe among institutions that are more sophisticated in the trading process (either owning independent trading desk or having larger trading volume) and among those 'discount' brokers who focus more on minimizing the trading cost of their clients.

4.3. Time-series Trend of the 'Round-number' Bias

Aimed at reducing both the latency and the cost of executing orders, especially those large orders from institutional investors, the algorithmic trading (AT) has grown rapidly since the late 1980s (often termed as 'program trading' in that time). Though the term 'AT' (or 'automated trading') is not a brand new concept, it has received ever-increasing attention in recent decades mostly due to the drastic development of the high-frequency trading (HFT) 'arena'. According to a recent report issued by a research team in Deutsche Bank, the share of the U.S. equity transaction via HFT reached to the peak at 60% in 2009 and declined gradually to about 50% in 2014[20]. As pointed out by Hendershott et al. (2011), the term 'AT' is quite broad and has a wide range of applications for different players in the market, whenever they apply computer algorithms in making portfolio decisions, or in submitting or cancelling orders[21]. One of the prominent features of the AT is the removal of human interventions, especially during the order submission and execution process (Foucault, 2012). For instance, based on a two-year trading data of a stratified sample of 120 stocks provided by the NASDAQ, Brogaard, Hendershott and Riordan (2014) provide a series of empirical evidences showing that liquidity demanding orders from high-frequency traders (HFTRs) are in the same of the permanent price change (implied by their state space model) and HFTRs' liquidity supplying orders are adversely selected. Hence, one would expect that with the widely spreading of the AT and HFT technologies among those active participants in the investment industry, the degree of the 'round-number' bias documented above shall be gradually reduced over time (H3). And we will formally test the causal effect of the AT on the market-level 'round-number' bias in Sec. 4.4, using the 2003 'auto-quote' event in NYSE as the exogenous shock.

[Insert Table 6 here]

In Table 6, we conduct 4 panel regressions based on our AN and TAQ sample, respectively. In order to estimate the magnitude of the time-series change of the 'round ratio' during our sample period, we intertwine our four dummy variables (i.e. D_00, D_50, D_X0, D_X5) with the Year variable (which counts the number of years from our sample beginning year, 2001). Therefore, one can compute the total reduction in terms of the average daily 'round ratio' in a later sample year comparing

---

[20] And they attribute the recent decline both in the market share and the profit of HFT companies to the relentless competition within the industry and the increasing cost of the infrastructures such as the co-location service and etc. See more detailed information at https://www.dbresearch.com/PROD/RPS_EN-PROD/PROD0000000000454703/Research_Briefing%3A_High-frequency_trading.pdf

[21] In 2010, the Australian Securities Exchange (ASX) issued a comprehensive report reviewing the status quo of the AT at ASX and the potential impact on the trading environment. In that report, the AT is defined as "computer generated trading activity whose parameters3 are determined by strict adherence to a pre-determined set of rules aimed at delivering specific execution outcomes." (https://www.asx.com.au/documents/media/20100211_review_algorithmic_trading_and_market_access.pdf)

with that in the starting year (2001) by simply multiplying the Year variable with the estimated coefficient. For instance, in our final year of the TAQ sample (2014), the average daily 'round ratio' for our sample TAQ trades executed at '.00' is only 1.56%, which is slightly above the unconditional expected value (1%) under the uniform distribution. And the annual reduction for different types of 'round number' prices is in line with the hierarchy documented in Section 4.1: the reduction is the largest for the 'integer' price '.00', followed by '.50', '.X0' and '.X5'. And this declining trend is also observed in our AN sample: (Column 3 & 4). In particular, the annual reduction of the 'round ratio (N)' at the integer price '.00' for an average institution (value-weighted across all sample institutions using the number of trades as the weight) is 0.2%, which is twice of the reduction speed in TAQ sample. Yet, for the 'round ratio (D)', the magnitude of the annual reduction in the TAQ sample (also VW across all sample stocks per day using the dollar amount of the trade as the weight) is larger than that of the AN sample. One possible explanation for this result is that given the fact that the TAQ captures the majority of the trading activities in the U.S. stock exchanges, the dollar-weighted 'round ratio' across all sample stocks is largely determined by the price preference of those large and active institutional investors in the market. And since the AN data set only covers a subset of the institutional investors in the U.S. (around 10% according to Puckett and Yan, 2011), the larger annual reduction in the 'round ratio (D)' within the TAQ sample suggests that institutional investors exhibit lesser degree of such behavioral bias across time. To sum up, the estimation results in Table 6 are consistent with our conjecture that with the rapid development in the AT practice, the degree of transaction prices clustered at those 'round-number' prices (e.g. '.50', '.10') would decline over time. And in the following section, we will use the 2003 'auto-quote' event occurred in the NYSE to offer some causal evidences regarding our 3rd hypothesis.

4.4 The 'Auto-quote' Event

In 2003, the NYSE began to implement the automated dissemination of its best bid and offer via SuperDOT limit orders within the exchange in a staggered fashion, which is referred as the 'auto-quote' event by Hendershott, Jones and Menkveld (2011). Briefly speaking, this initiative required that the NYSE would 'autoquote' the highest bid or lowest offer whenever there was a relevant change to a specialist's limit order book. However, this amended rule still allows specialists to submit and disseminate quotes manually on behalf of their own trading interests or their floor traders. Hence, as pointed out by Hendershott et al. (2011), this technical improvement in NYSE helped to reduce the labor work of clerks at the specialist posts in terms of order matching and quote updating and the latter one is crucial for algorithmic traders in order to be better informed about the current effective quote and submit marketable limit orders via the NYSE's automatic execution facility (then named as *NYSE Direct+*®). In the meantime, the implementation of this 'auto-quote' system would not affect the trading behavior of human traders as they only witness the reduction in quote updating and still reply on the conventional floor traders or the new system to execute their trades.

Following their study, we define the pre-event period as two months before January 29, 2003, the starting date of the implementation and the post-event period as two months after the date of the last adoption, which is May 27, 2003. As this event only occurred at the NYSE, in order to quantify the impact of this staggered adoption of the 'auto-quote' system on the degree of the 'round-number' bias in the overall market, here we also include the NASDAQ-listed stocks from our TAQ sample in the same period as the control group and conduct the pair-matching based on the industry and book-to-market ratio as of the closest fiscal year (as in Chordia and Miao, 2019). Hence, in this section, we carry out a 'difference-in-difference' empirical design and report the estimation results in Table 7 & 8. The main variable of interest is the daily stock-level 'round ratio (N)' in Table 7 and in Table 8 it is the daily stock-level 'effective cost' at different types of 'rounded penny' prices (value-weighted by the dollar volume within the same stock per day). And we conduct both the univariate and regression analysis to estimate the economic impact of this adoption. In Panel A of Table 7, the decline of the average 'round ratio' within the NYSE stocks after the event is the largest for the whole number, '.00', followed by '.50', '.X0', '.X5'. And the drop is only statistically significant among stocks listed in NYSE, not for NASDAQ stocks. In Panel B, we include some exchange-level control variables and estimate the marginal difference of the 'round ratio' of the NYSE stocks (value-weighted across all sample stocks per day) comparing with those listed in NASDAQ before and after the 'auto-quote' event. Not only the reduction is statistically significant during the two-month window after the adoption, the economic magnitude is also in accordance with that in the Panel A of Table 7. Here, to our surprise, the deduction is most significant for the 'half-dollar' price '.50' (from 1.83% per day two months before the event to 1.73% per day two months afterwards, while those paired NASDAQ stocks experienced a significant increase at the 'half-dollar' price after the event, by around 0.19%), followed by the drop in 'integer', 'dimes' and 'nickels', as in Table 8. To sum up, due to the installment of this automated order routing and disseminating system in 2003, we witness a significant drop in the 'round ratio' of all four types of 'round-number' prices. Combining with the evidence in Hendershott et al. (2011) that the electronic message per minute ($100 trading volume) increases after the staggered adoption of the system, evidences in Table 6 suggest that the increased use of algorithmic trading systems has led to the reduction in the percentage of trades executed at 'round-number' prices. In Section 4.5, we will present estimations of the benefit due to the reduction in the 'round-number' bias.

4.5. The Economic Consequence of the 'Round-number' Bias and the Estimated Welfare Loss

In order to formally test our 4th hypothesis, which conjectures that due to the overly crowding of trades executed at certain 'round-number' prices the execution cost will be pushed up, following the common practice in microstructure literature, we compute the value-weighted 'effective cost' measure for our sample trades executed at different types of 'round-number' prices. As in Hasbrouck (2009), we match each sample trade with contemporaneous effective NBBO quote and define the 'effective cost' as follows:

$effective\ cost_i = trade\ direction_i * [\log(price_i) - \log(midpoint_i)]$ Eq (3),

where $trade\ direction_i$ equals 1(-1) if the trade i is classified as a 'buyer-initiated' ('seller-initiated') trade based on the Lee-Ready algorithm. Then, within a trading day, for a certain 'rounded penny' price (say '.50'), we compute the value-weighted (VW) 'effective cost' for the same sample stock based on the dollar trading volume of each trade. Next, we can compute two versions of 'effective cost' for a 'rounded penny' price, by aggregating across all sample stocks in a trading day in an equal-weighted (EW) or a VW (using the dollar trading volume of the stock as the weight) fashion. Finally, we perform an OLS regression each day with the dependent variable being the VW/EW 'effective cost' of a certain 'rounded penny' price (in total 100 observations per day), similar to the Equation (1).

[Insert the Table 9 here]

The estimation results reported in Table 9 indeed confirm our last hypothesis that due to the abnormal clustering of trades which are executed at or around 'round-number' prices ('.00', '.50', '.X0' and '.X5'), the transaction cost (comparing with the midpoint of the NBBO quote) is relatively higher comparing with that of the 'non-round-number' price. In terms of the economic magnitude, we focus on the coefficient estimated based on the value-weighted method, as under this approach, we can directly estimate the economic impact of such 'round-number' bias during the trading process. Within our sample period for the TAQ sample (2001-2014), the total dollar trading volume amounts to 350 trillion dollars among which 11 trillion (3.2%) are executed at the 'integer' price ('.00'), followed by 'half-dollar' (2.1%), 'dimes' ('.X0', 12%) and 'nickels' ('.X5', 13%) (See Table 8). Hence, one can get a rough estimation of the additional trading cost that investors have spent on trades executed at '.00' or '.50' by multiplying the estimated coefficient (in Column 1) with the corresponding trading volume (see the detailed estimation below).

Following prior empirical microstructure studies that analyze the trading cost of certain strategies (e.g. Lesmond, Schill and Zhou, 2004; Novy-Marx and Velikov, 2015), here we mainly use the 'effective cost' as the measurement for our trade-level data[22]. As discussed in our hypothesis development in Chapter 2, if such price preference exists in the market-wide, then both the explicit and implicit cost of executing an order would be affected. For instance, using the historical order submission and execution data of the Taiwan Futures Exchange, Kuo et al. (2015) report that those limit orders with prices targeting the 'round-number' end up with lower execution ratio and take longer time to be executed. Hence, combining with those models on the bidding and the market making process (e.g. Foucault, Kadan and Kandel, 2005) where impatient traders pay premiums to get in the front of the limit order book, we conjecture that orders that are executed at such 'round-number' prices shall on average be more costly comparing with those executed at 'non-round-number' prices (e.g. '.X1' where X is an integer ranging from 1 to 9).

Specifically, we run a daily cross-sectional regression using Equation (1) but replacing the 'round ratio' with the VW 'effective cost' of trades executed at certain type of 'rounded penny' price(s). Then we compute the time-series average of the

---

[22] In Bhattacharya, Holden and Jacobsen (2012), they estimate the 'wealth transfer' caused by investors' preference for submitting buy (sell) orders which are one penny less (higher) than a round-number price, e.g. buying at 'X.99' and selling at 'X.01', based on the hypothetical 24-hour return of each trade. Yet, their intention is trying to measure the temporary buying or selling pressure caused by such behavioral bias, not the trading cost per se. Hence, in this study, we choose the 'effective cost (spread)' as the proxy for direct trading cost, as it is most reliable and widely used in microstructure studies (see the detailed discussion in Goyenko, Holden and Trzcinka (2009)).

coefficient within each calendar year. The reason that we perform such analysis annually is the dramatic time-series change of the in-sample ‘effective cost’ as documented in Section 4.4 and we want to pose a more informative picture of the dynamics of such bias in the first decade of the 21th century[23]. From the Panel A of Table 10, one can tell that in the first 5 years of the TAQ sample, the percentage ‘effective cost’ of trades executed at the ‘whole number’ price (‘.00’), the ‘half-dollar’ (‘.50’), the ‘dimes’ (‘.X0’), or the ‘nickels’ (‘.X5’), is much larger than that of a trade executed at any of ‘non-round-number’ prices (e.g. ‘.13’, ‘.17’ and etc.). And this gap has declined to less than 1 bps approaching to the end of our sample period. To offer some benchmark numbers of our estimated ‘effective cost’, one study done by Hasbrouck (2009) shows that the median of the ‘effective cost’ estimated based on a sample of 300 stratified stocks has declined to around 5 to 6 bps after 2004. Another study done by Anand et al. (2013) presents some evidences showing that the average ‘effective cost’ of the TAQ trade has been reduced to about 4 bps in 2011 yet peaked to around 10 bps in the midst of 2008 financial crisis.

[Insert Table 10 here]

Next, we aggregate the total dollar trading volume of the sample TAQ trades executed at each ‘rounded penny’ price (e.g. ‘.01’, ‘.02’ and etc.) each calendar year (Panel B). And finally, we estimate the additional transaction cost (measured by the ‘effective cost’ in dollar term) in two steps: first we compute the gap of the ‘effective cost’ between trades executed at certain ‘round number’ price(s) and those executed at the remaining ‘non-round-number’ prices (i.e. ‘Any other price’ in Panel A); then, we multiply this gap with the dollar amount of trades executed at that type of ‘round number’ price(s) in the same calendar year and derive the estimated ‘welfare loss’ in Panel C. Similar to the pattern of the ‘effective cost’ in Panel A, investors paid a significant proportion of this additional trading cost in the first 5 years of our sample period. Notably in 2007 and 2008, there was a slight increase in the dollar amount of trades executed at the ‘whole dollar’ price which led to the increase in the trading cost.

---

[23] And the reason that we choose the ‘effective cost’ as our main measure of trading cost is based on a comprehensive study done by Goyenko, Holden and Trzcinka (2009), in which they examine both the cross-sectional and time-series correlations of some prominent low-frequency liquidity measures (e.g. the Amihud’s Illiquidity) with measures from the high-frequency data set (e.g. the TAQ). And the benchmark measures in their study are mainly the effective spread and the five-minute price impact measure proposed by Hasbrouck (2009). Since the latter is both determined by the aggressiveness of the bidding order and the size of the order submitted, the most relevant trading cost measure is the ‘effective cost’ in our setting. In addition, it can be viewed as the lower bound estimate of the economic consequence of the ‘round-number’ bias in the domain of security trading.

Nevertheless, within our TAQ sample (2001-2014), under the value-weight scheme (by dollar amount), trades executed at rounded ‘whole dollar’ (‘half-dollar’) prices cost around 1.7 (1.8) bps more than those executed at ‘non-round-number’ prices (e.g. ‘.13’, ‘.17’ and etc.). And they contribute to about 13% (12%) of the additional trading cost (measured by the ‘effective cost’ in dollar term) caused by the overly crowded trades at those ‘round-number’ prices. In total, this kind of behavioral bias has cost the U.S. investors, both institutions and retail types, about 11 billion of dollars within our sample period. Yet, one can also tell that the majority of this cost concentrates in the first two years (2001-2002) of our TAQ sample, which is worthy of exploring in future studies. To help readers better interpret our estimations, we attach the ‘wealth transfer’ estimation made by Bhattacharya et al. (2012). In that study, the authors focus on the temporal price pressure caused by the ‘left-digit’ effect that leads to the increased buying (or selling) orders when the ask falls (or the bid increases) to reach the ‘integer’ price (e.g. ‘.10’, ‘.15’). Hence they use the 24-hour return measure by assuming the buy (or short-sale) position is closed 24 hours later (also assuming the stock is quite liquid in this short horizon). Their estimations show that within their sample period (2001 to 2006) those liquidity demanders whose orders are clustered one penny above or below those integer/half-dollar prices/‘dimes’/ ‘nickels’ on average lose around 600 to 1,000 million to their counterparties per year. And our annual estimate of this additional explicit trading cost ranges from around 200 to 900 million of dollars when the first two years are excluded. One caveat of their ‘wealth transfer’ estimation method is rooted from their assumption: though their ‘one-day’ inventory window eliminates the effect of other informational events, investors might be better off by holding on their ‘biased’ positions longer. We believe our estimation is also quite conservative as we mainly focus on the ‘effective cost’ measure, which directly comparing the ‘biased’ execution price with the mid-point of the NBBO quote.

Next, in order to investigate the impact of the price clustering around ‘round-number’ prices for the institutional investors, following some prior studies (Puckett and Yan, 2011; Anand et al., 2012), we conduct both univariate sorting analysis as well as the cross-sectional regression analysis in order to test whether institutions’ ‘round-number’ bias is related to their trading performance. Building on the current consensus of the measurement of institutions’ trading costs, we adopt the following method as the proxy for the ex-post measure of transaction costs incurred by our sample institutions:

$Trading\ cost_{i,k,t} = order\ direction_{i,k,t} * \left(price_{i,k,t} - benchmark\ price_{i,k,t}\right)/ benchmark\ price_{i,k,t}\ \ Eq\ (4)$,

where $price_{i,k,t}$ is the execution price of the k-th trade submitted by the institution i in day t and the $benchmark\ price_{i,k,t}$ is the appropriate benchmark price of this trade in day t and the $order\ direction_{i,k,t}$ equals 1 for the buy order and -1 for the sell order. As noted by Hu (2009), the widely used pre- and post-trade measures (using the closing price one day before or on the execution day, respectively) suffer from market movements which may dampen the interpretation of the result. Hence, here we rely on the 'during-trade' measure by using the volume-weighted average price (VWAP) as the 'benchmark price' (Berkowitz, Logue and Noser, 1988). For each institution, we compute the dollar-weighted percentage trading cost measure of all sample trades executed within a calendar quarter and term it as 'TC_VWAP'.

In the common practice, institutional investors disclose reports of their portfolio holding and performance on a quarterly basis, hence, following Puckett and Yan (2011) and Kuo et al. (2015), we measure the degree of 'round-number' clustering and the percentage trading cost of our sample institutions at quarterly frequency. Specifically, for each quarter, we sort our sample AN institutions (with unique 'clientcode') into quintiles based on their total percentage of trades executed around '.X0' and '.X5' prices and compute the equal-weighted average of the quarterly 'TC_VWAP' within each quintile. Then we conduct the Student's t-test on the difference of the average of the 'TC_VWAP' of those sample institutions sorted in the top and bottom quintile, up to 3 quarters forward. Based on the results shown in the Panel A of Table A4, there is clear pattern of negative correlation between the 'round-number' clustering and the dollar-weighted trading cost measure within our sample institutions. For instance, in a calendar quarter, those institutions with on average 32% of trades executed around 'dimes' and 'nickels' prices within the same quarter, end up with only 2 basis points VWAP-based trading costs; yet those institutions whose 'round ratios' are close to the expected value under the uniform distribution (20%) incur over 6 basis points under the VWAP measurement. And the statistical significance of this 'trading skill' gap lasts till 3 quarters ahead. Furthermore, we also conduct the multivariate regression analysis in Table A5, by including several control variables related to the characters of the size of the institutional trading as well as the stocks traded by the institution. For instance, an early theoretical study done by Saar (2001) argues that the degree of price 'run-up' (i.e. 'momentum') can have significant impact on the market impact of institutions' buy and sell orders. Hence, we include 3 variables related to the characters of stocks traded by a sample institution, namely the average price 'momentum' (7-month prior to the current month), the average daily turnover ratio and the mean of the logarithm market capitalization (in millions) of the stocks traded by institution in the current quarter. Due to the anonymity of the AN data set, here we only include the logarithm of the total number of trades in the previous quarter, in order to control for the difference in the size and frequency of the institution's inter-quarter trading activities[24]. Based on the estimation results, one can still find a significant negative relation between the quarterly percentage trading cost and the 'round ratio' in the last quarter (and the persistence of the quarterly 'round ratio' is quite strong, as shown in Panel B of Table 4A).

To interpret this peculiar finding, we offer a possible explanation based on some prior literature on the trading skill and particular trading strategies utilized by those sophisticated institutional investors. To name a few, by using the same data set, Anand et al. (2012) reports significant negative 'execution shortfall' (on monthly basis, see their Table 2 in page 568) for a subset of AN sample institutions and their superior trading skill can last at least 4 months ahead. Since by subscribing the trading costs analysis of Abel Noser, these sample institutions indicated their special needs on monitoring their trading costs and their average size is relatively larger comparing with an average 13-F institution (Puckett and Yan, 2011), we believe that some of them are able to minimize or even earn additional profits by acting as liquidity providers strategically. Indeed, a recent study done by Jame (2018) which focuses on a subset of identifiable hedge funds in the AN data presents evidences showing that some of them are able to earn additional profits by acting as 'counterparties' for those mutual funds with higher funding constraints. Hence, one plausible cause for this negative relation between institutions' trading cost and their 'round ratio' is that they are able to act strategically by consuming liquidity (submitting market orders) when the order-book around a 'round-number' price is thick and submitting limit orders

[24] In addition, in order to control for the unobserved factors that are related to the trading skill of our sample institutions, here we conduct the panel regression with fixed-effect at 'clientcode' level and include dummy variables for each quarter.

when the liquidity becomes scarce around a certain ‘round-number’ price. Yet, due to that we cannot differentiate the order type in the AN data set, we leave this unexplored yet quite interesting pattern to future research.

Chapter 5. Conclusion and Directions for Future Research

In this study, I revisit a prevalent cognitive pattern that is inherent to the human brain: people rely on ‘round-numbers’ (e.g. integers, half, quarter) to save their cognitive energy and speed up the interpretation process, in the context of the daily trading activities in the U.S. stock market. First, based on the TAQ sample from 2001 (post-decimalization) to 2014, I document a strong pattern that the execution price of the sample trade is overly clustered at a series of ‘round-number’ prices. And the degree of the clustering follows a hierarchy which is consistent with our daily experience: ‘.00’, ‘.50’, quarters, dimes and nickels, which I term as ‘round-number’ bias. Prior studies have offered some rationales for such price-clustering phenomenon, e.g. the valuation uncertainty/negotiation hypothesis (Ball, Torous and Tschoegl, 1985; Harris, 1991) and the collusion argument made by Christie and Schultz (1994). Yet, as pointed out by Ikenberry and Weston (2007), the magnitude of such clustering is too large to be explained by those empirical proxies of valuation uncertainty and market structure differences. Indeed, I find that though the degree of such ‘round-number’ bias is decreasing in time, at the end of our sample period (2014), the distribution of the transaction price over 100 penny intervals is still different from the uniform distribution.

In order to further lend support to our conjecture that this clustering pattern is rooted from human investors’ cognitive limitation, I move on to a transaction-level data set of some institutional investment firms in the U.S., provided by Abel Noser. First, I find that though being viewed as ‘sophisticated traders’ in the market, those sample institutions also exhibit a significant preference for ‘round-number’ prices. Yet, when comparing with the TAQ sample on a daily basis, the degree of the ‘round-number’ bias in the AN sample is much more smaller, which suggests that those retail investors contribute more to the overall degree of price clustering in the market. Moreover, I make full use of the AN data by exploring the heterogeneities across different types of institutions and brokers in the sample and find that those institutions who hire execution traders (or have larger trading volume) and those ‘execution-only’ brokers who charge less per share commissions exhibit lesser degree of such bias in their submitted or executed trades. Finally, based on the percentage ‘effective spread’ measurement, I make a conservative estimation that due to the overly clustering on the ‘round-number’ prices, investors in the U.S. stock market on average incur additional transaction costs which amount to 785 million annually during our sample period (2001-2014).

One caveat of our study is that I did not provide direct evidence on the degree of the ‘round-number’ bias among trades submitted by retail investors, as I do not have a good sample of retail trades in the U.S. after decimalization and according to a recent study by Boehmer et al. (2019) those retail-initiated trades tend to receive price improvements. Another potential future exploration worthy of investigating is whether rounding is beneficial in some unusual settings, e.g. the Flash-Crash event in 2010, and whether certain types of traders (e.g. the hedge fund or securities brokers) take advantage of investors’ cognitive limitation by strategically submit limit or market orders in certain circumstances.

Tables and Figures for Chapter 4

Table 1. Summary Statistics

This table reports the summary statistics of the TAQ sample trades from Feb. 2001 to Dec. 2014 (Panel A) and that of the Abel Noser sample institutional trades from Feb. 2001 to Sep. 2011 (Panel B). For each sample, we first round the transaction price to the nearest penny. The sample selection criteria for the TAQ trade are in line with Chordia et al. (2001). We match with the PERMNO from the CRSP stock database with SHRCD equal to 10 or 11. And we only include trades submitted by institutions with client type equal to 1 or 2: pension plan sponsors or investment managers. Following Hasbrouck (2009), the 'effective cost' measure is defined as the difference between the log transaction price and the log mid-point quote, multiplied by (-1) if it is a sell trade. We apply the Lee & Ready (2001) algorithm to determine the direction of each sample trade. And we compute the daily VW 'effective cost' measure based on all sample trades executed at a given 'rounded penny' price, using the dollar trading volume as the weight. We report the time-series average of the daily average of each statistics within the same calendar year.

Panel A: Descriptive statistics of the TAQ sample trade

Year	Annual Number of Common Stocks	Average Daily No. of trades (in millions)	Daily VW 'effective cost'	Daily Trading Volume (in millions)	Average Trade Size (in dollars)
2001	5652	3.30	0.12%	89,405	27,091
2002	4890	3.77	0.08%	79,447	21,063
2003	4462	4.65	0.06%	76,996	16,555
2004	4473	6.09	0.05%	95,892	15,746
2005	4618	7.53	0.04%	116,601	15,493
2006	4525	10.40	0.04%	147,303	14,168
2007	4469	18.44	0.04%	234,200	12,703
2008	4307	33.99	0.05%	303,157	8,918
2009	3578	34.04	0.05%	234,097	6,878
2010	3359	28.68	0.04%	301,826	10,523
2011	3453	30.22	0.04%	248,276	8,215
2012	3208	24.59	0.04%	206,271	8,390
2013	2977	23.51	0.05%	218,813	9,307
2014	2830	29.91	0.04%	258,597	8,645

Panel B: Descriptive statistics of the Abel Noser sample trade

Year	Average Daily No. of trades	Daily No. of institutions	Daily Number of brokers	Daily Trading Volume (in millions)	Average Trade Size (in dollars)
2001	27,779	394	712	8,812	317,226
2002	38,717	424	781	10,153	262,246
2003	38,291	400	767	8,787	229,469
2004	45,983	402	747	9,583	208,408
2005	62,774	376	809	11,552	184,029
2006	107,181	399	790	14,679	136,953
2007	136,062	377	791	15,867	116,615
2008	113,097	334	770	15,088	133,405
2009	91,165	317	720	10,797	118,432
2010	92,635	307	717	9,893	106,794
2011	91,339	258	617	8,939	97,871

Table 2: Tests for the frequency distribution of trades executed at different 'rounded penny' prices

Panel A: The Abel Noser sample (2001-2011)

H_0: $Freq(".XY") = 0.01$, where $X, Y\ in\ (0, 1, \ldots, 9)$

Statistic tests	**Testing statistic**	**p-value**
Chi-square test	2670.67	<.0001
Likelihood-ratio test	2278.39	<.0001
Kolmogorov-Smirnov test	0.9758	<0.001
Anderson-Darling	356.09	<0.001

Panel B: The TAQ sample (2001-2014)

H_0: $Freq(".XY") = 0.01$, where $X, Y\ in\ (0, 1, \ldots, 9)$

Statistic tests	**Testing statistic**	**p-value**
Chi-square test	268401.9142	<.0001
Likelihood-ratio test	236308.1464	<.0001
Kolmogorov-Smirnov test	0.9786	<0.001
Anderson-Darling	355.7468	<0.001

Panel C: Statistical tests for the difference of the frequency distribution: Abel Noser vs. TAQ sample (2001-2011)

H_0: $EDF(".XY"|Abel\ Noser\ trades) = EDF(".\mathrm{XY}"|TAQ\ trades)$, where $X, Y\ in\ (0, 1, \ldots, 9)$

Statistic tests	**Testing statistic**	**p-value**
Kolmogorov-Smirnov test	1.2728	0.0783
Kuiper test	1.9800	0.0116
F-test	144.59	<0.001

* Here, in Panel A and B, we report the statistic tests for the null hypothesis of uniform distribution of transactions executed at different 'round-number' prices (e.g. 'X.01', 'X.02', ... , 'X.99') of the sample trade in Abel Noser and of the TAQ sample. And in Panel C, we perform the statistical tests that compare the empirical cumulative distribution function (EDF) of the total number of trades that are executed at a 'round-number' price below ".XY", where X, Y can change from 0 to 9, between the Abel Noser and TAQ sample. Here, we only include trades that can be matched to the 'PERMNO' code from the CRSP database with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3 and submitted by institutions with client type being 1 or 2, i.e. plan sponsors or investment managers.

For each sample transaction, we first round the transaction price to the penny level. Then, for each sample stock, we calculate the total number of transactions executed at a 'rounded penny' price, e.g. 'X.20', and then scale by the total number of transactions executed for the same stock at the same trading day. Next, we compute the weighted average of this ratio across all sample stocks within the same trading day, using the number of trades as the weight, for a given 'rounded penny' price. Finally we compute the arithmetic mean of this ratio across all sample trading days and test whether it follows the uniform distribution. Under the null hypothesis (i.e. no price clustering), this ratio should be 0.01 and sums up to 1 during a trading day. The Pearson's Chi-square (also the Likelihood-ratio) test in Panel A and B can be applied for the discrete distribution and the Kolmogorov-Smirnov (and the Anderson-Darling) test Panel A and B suits for the continuous setting. And in Panel C, the implicit assumption of the Kolmogorov-Smirnov (Kuiper) test is that the underlying distribution function of a variable is continuous. In addition, we provide the result based on the F-test as well, which is designed for discrete observations that are binned at certain frequency (here $0.01), as used by Ikenberry and Weston (2007).

Table 3: Tests for the frequency distribution of trades executed at different 'rounded penny' prices across subsamples in Abel Noser (2001-2011)

Statistic tests	Testing statistic	p-value
Panel A: Statistical tests for the difference of the frequency distribution: trades with vs. without 'valid' 'trader' codes (the Abel Noser sample, 2001-2011)		
$H_0: EDF(".XY"\|'Trader') = EDF(".XY"\|Non\ 'Trader')$, where $X, Y\ in\ (0, 1, \ldots, 9)$		
Statistic tests	**Testing statistic**	**p-value**
Kolmogorov-Smirnov test	0.6364	0.8127
Kuiper test	0.9899	0.8342
F-test	1.57	0.0129
Panel B: Statistical tests for the difference of the frequency distribution: trades executed by 'discount' vs. 'full-service' brokers (the Abel Noser sample, 2001-2011)		
$H_0: EDF(".XY"\|'discount'\ brokers) = EDF(".XY"\|'full-service'brokers)$, where $X, Y\ in\ (0, 1, \ldots, 9)$		
Statistic tests	**Testing statistic**	**p-value**
Kolmogorov-Smirnov test	1.2727	0.0783
Kuiper test	1.6970	0.0663
F-test	1.57	0.0129
Panel C: Statistical tests for the difference of the frequency distribution: 'Big' vs. 'Small' institutions (the Abel Noser sample, 2001-2011)		
$H_0: EDF(.XY\|'Big'institutions) = EDF(.XY\|'Small'institutions)$, where $X, Y\ in\ (0, 1, \ldots, 9)$		
Statistic tests	**Testing statistic**	**p-value**
Kolmogorov-Smirnov test	1.0606	0.2106
Kuiper test	1.3435	0.3366
F-test	75.75	<.0001

* Here, in Panel A, we perform the statistical tests that compare the empirical cumulative distribution function (EDF) of the total number of trades that are executed at a 'round-number' price below ".XY", where X, Y can change from 0 to 9, between trades that can be matched with 'valid' trader codes and those without 'valid' trader codes. Here, 'valid' trader codes are those sample AN trades that can be matched with a 'tradercode' (not being purely numeric or missing) by the 'clienttdrcode' via the 'TraderXref' table provided by Abel Noser. And the sample AN trades are those that can be matched to the 'PERMNO' code from the CRSP database with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3 and submitted by institutions with client type being 1 or 2, i.e. plan sponsors or investment managers. In Panel B, we split our AN sample trades into two groups, based on whether the trade is executed by a broker ('brokercode') whose dollar-weighted commission per share is above or no more than 3 cents. In Panel C, those sample institutions (within unique 'clientcode') whose total dollar trading volume are ranked in the top (bottom) 33% are labelled as 'Big' ('Small') institutions within our Abel Noser sample.

For each sample transaction, we first round the transaction price to the penny level. Then, for each sample stock, we calculate the total number of transactions executed at a 'rounded penny' price, e.g. 'X.20', and then scale by the total number of transactions executed for the same stock at the same trading day. Next, we compute the weighted average of this ratio across all sample stocks within the same trading day, using the number of trades as the weight, for a given 'rounded penny' price. Finally we compute the arithmetic mean of this ratio across all sample trading days and test whether it follows the uniform distribution. Under the null hypothesis (i.e. no price clustering), this ratio should be 0.01 and sums up to 1 during a trading day. The Pearson's Chi-square (also the Likelihood-ratio) test in Panel A and B can be applied for the discrete distribution and the Kolmogorov-Smirnov (and the Anderson-Darling) test Panel A and B suits for the continuous setting. And in Panel C, the implicit assumption of the Kolmogorov-Smirnov (Kuiper) test is that the underlying distribution function of a variable is continuous. In addition, we provide the result based on the F-test as well, which is designed for discrete observations that are binned at certain frequency (here $0.01), as used by Ikenberry and Weston (2007).

Figure 1. 'Round-Number' Bias: All Abel Noser Sample Trades (2001-2011)

Following KLZ (2015), we first round the transaction price of all Abel Noser sample trades to the nearest penny, e.g. 'X.01', 'X.02' and etc. Then we aggregate them by number/shares/dollars at a given 'rounded price' (e.g. 'X.01') and scale by the total number/shares/dollars of trades conducted by all sample institutions within a trading day. And finally we calculate the arithmetic mean across all sample trading days (Feb.2001-Sep.2011) for a given 'rounded price' and plot accordingly. Here, we only include trades that can be matched to the 'PERMNO' code from the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3. And we only count trades submitted by institutions with client type being 1 or 2: plan sponsors or investment managers, respectively. Fig. 2A reports the time-series average of the frequency based on the number of trades executed at different 'round-number' prices while Fig. 1B and 1C report the figures based on dollars and shares traded, respectively.

Fig. 1A: 'Round-number bias' (based on numbers of trades)

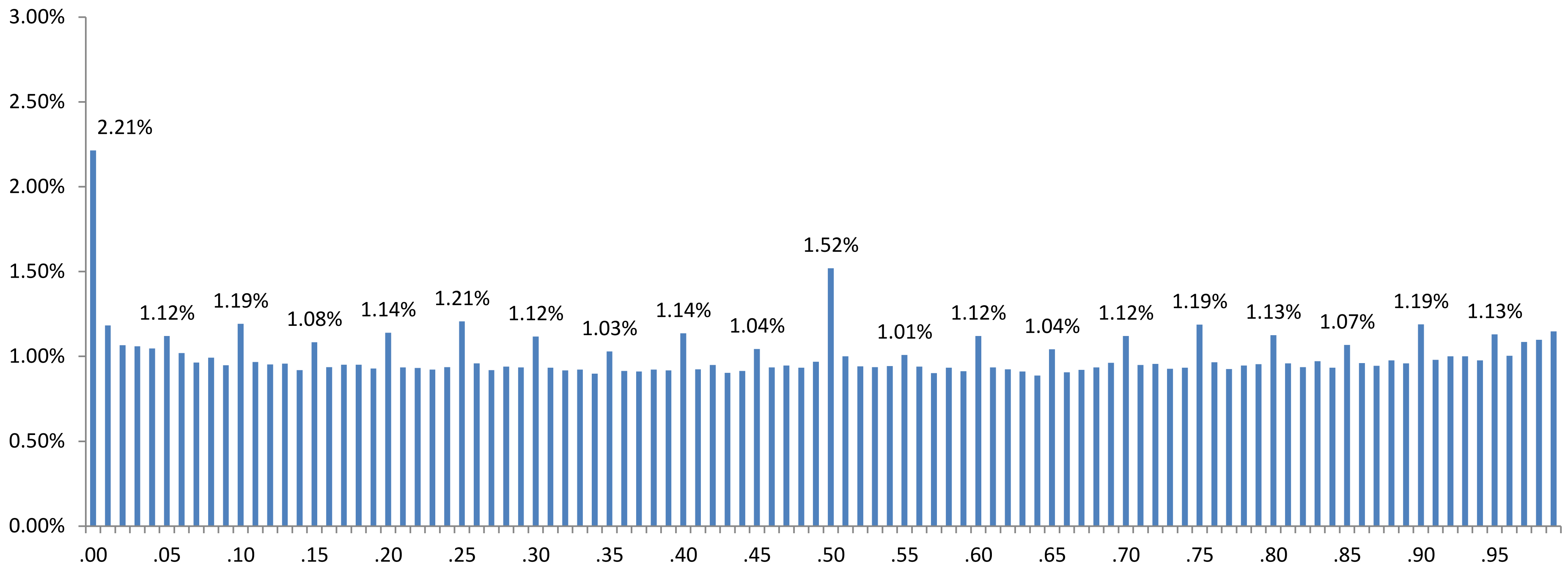


Fig. 1B: 'Round-number bias' (based on dollars traded)

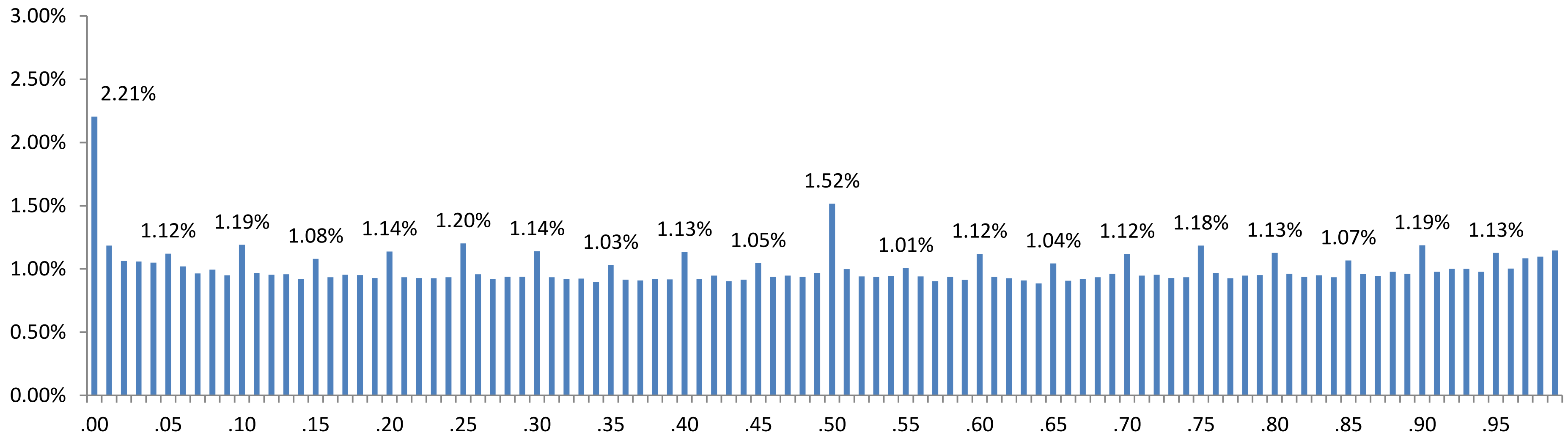


Fig. 1C: 'Round-number bias' (based on shares traded)

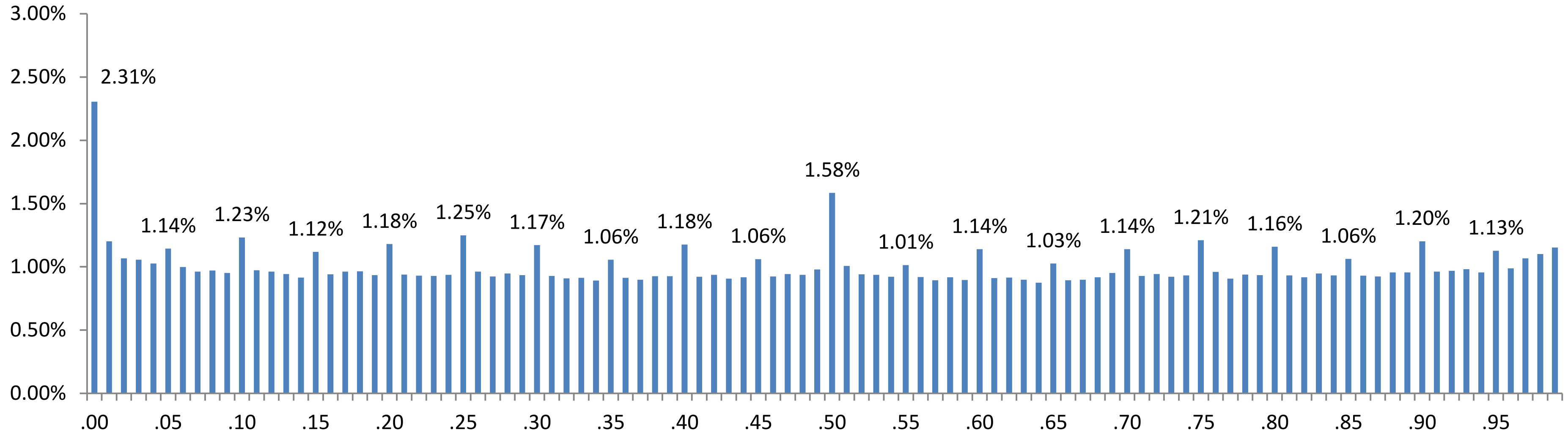


Figure 2: 'Round-Number' Bias: TAQ Sample Transactions (2001-2014)

Following KLZ (2015), we first round the transaction price of all sample TAQ trades to the nearest penny, e.g. 'X.01', 'X.02' and etc. Then we aggregate them by number/shares/dollars at a given 'rounded price' (e.g. 'X.01') and scale by the total number/shares/dollars of trades within a trading day. And finally we calculate the arithmetic mean across all sample trading days (Feb.2001-Dec.2014) for a given 'rounded price' and plot accordingly. Here, we only include trades that can be matched to the 'PERMNO' code in the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3 within the normal trading hour, i.e. from 9:30 a.m. to 4:00 p.m. in Eastern Time. Fig. 1A reports the time-series average of the frequency based on the number of trades executed at different 'round-number' prices while Fig. 1B and 1C report the figures based on dollars and shares traded, respectively.

Fig. 2A: 'Round-number bias' of TAQ trades (based on numbers of trades)

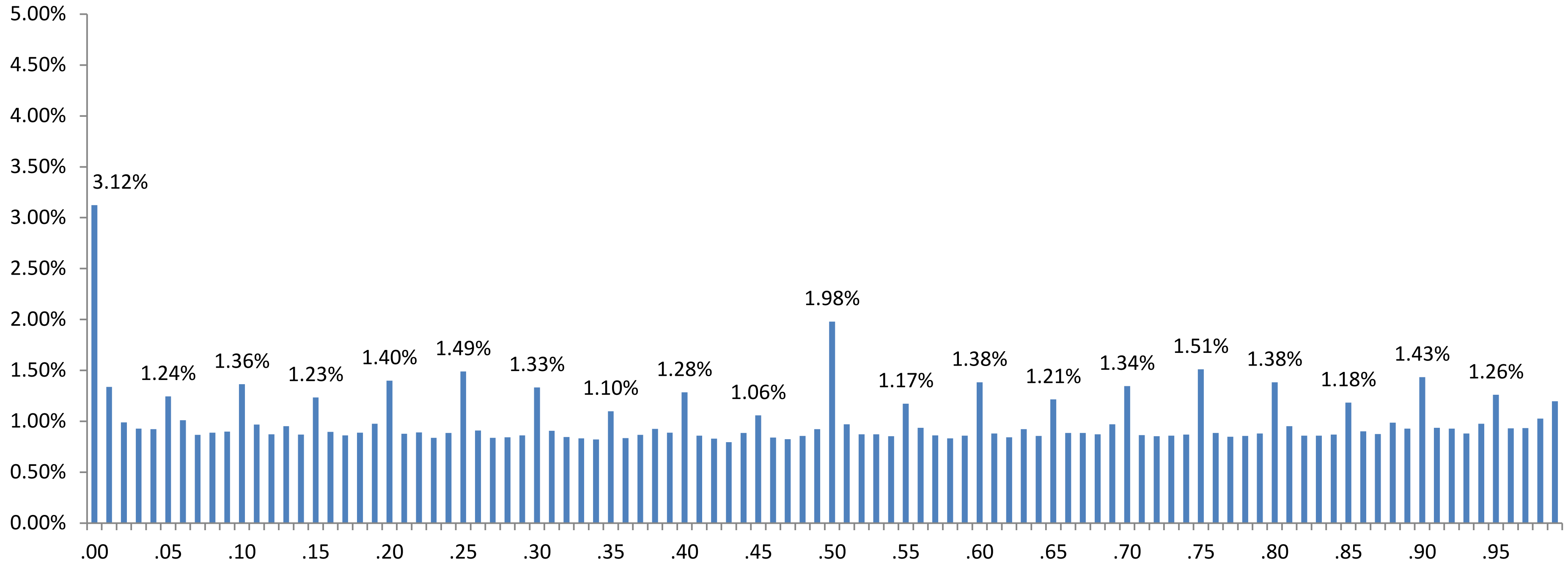


Fig. 2B: 'Round-number bias' of TAQ trades (based on dollars traded)

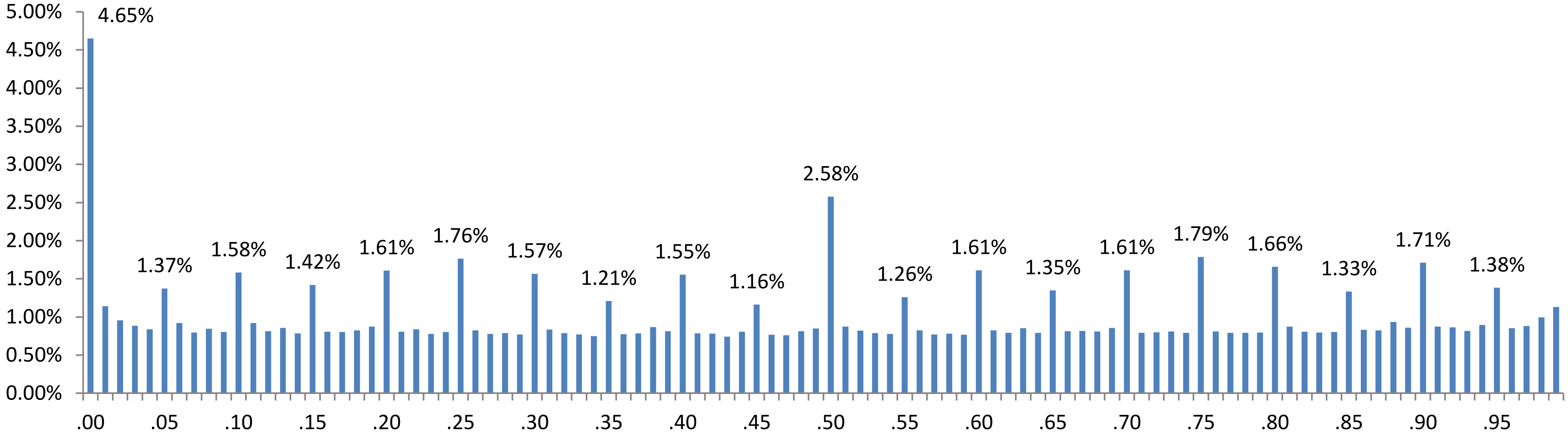


Fig. 2C: 'Round-number bias' of TAQ trades (based on shares traded)

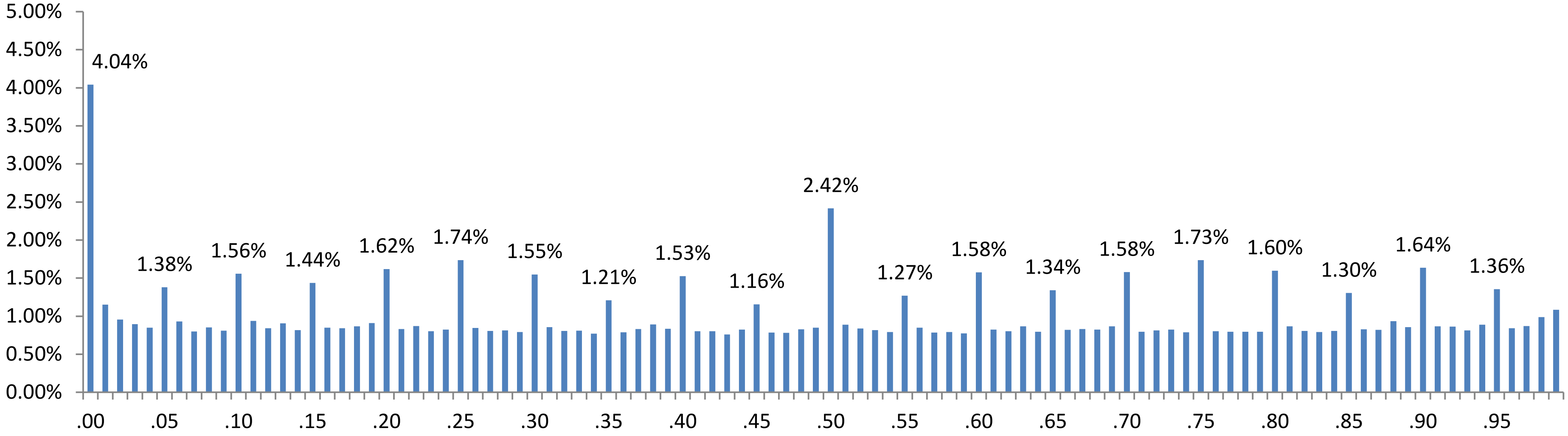


Figure 3: 'Round-Number' Bias: 'Traders' vs 'Non-Traders' (AN sample, 2001-2011)

Following KLZ (2015), we first round the transaction price of all 'valid' sub-sample trades to the nearest penny, e.g. 'X.01', 'X.02' and etc. Then we aggregate them by number at a given 'rounded price' (e.g. 'X.01') and scale by the total number of trades conducted by all sample institutions within a trading day to compute the 'round ratio'. And finally we calculate the average across all sample trading days (Feb.2001-Sep.2011) for a given 'rounded price' and plot accordingly. Here, 'valid' trades refer to the sample Abel Noser transactions that can be matched to the 'PERMNO' code from the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3. And we only included trades submitted by institutions with client type being 1 or 2, i.e. plan sponsors and investment managers. In Fig. 3A, we report the average daily 'round ratio' within the subsample of trades that can be matched with a valid 'tradercode' via the 'TraderXref' table provided by Abel Noser and without being purely numeric, e.g. '384', '280'. Fig. 3B reports the average daily 'round ratio' within the subsample of Abel Noser trades that cannot be matched with any valid 'tradercode'. The sample starts from Feb. 2001 to Sep. 2011.

Fig. 3A: 'Round-number bias' of 'traders' (based on the number of trades)

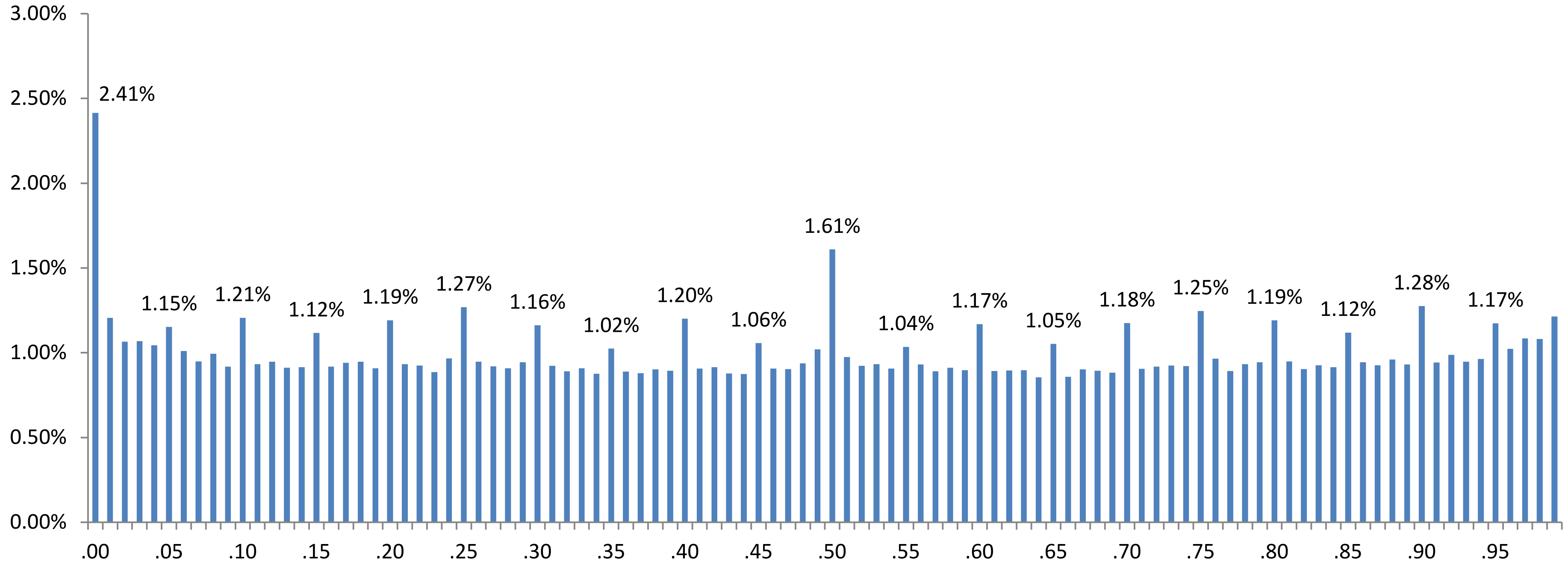


Fig. 3B: 'Round-number bias' of non-'traders' (based on the number of trades)

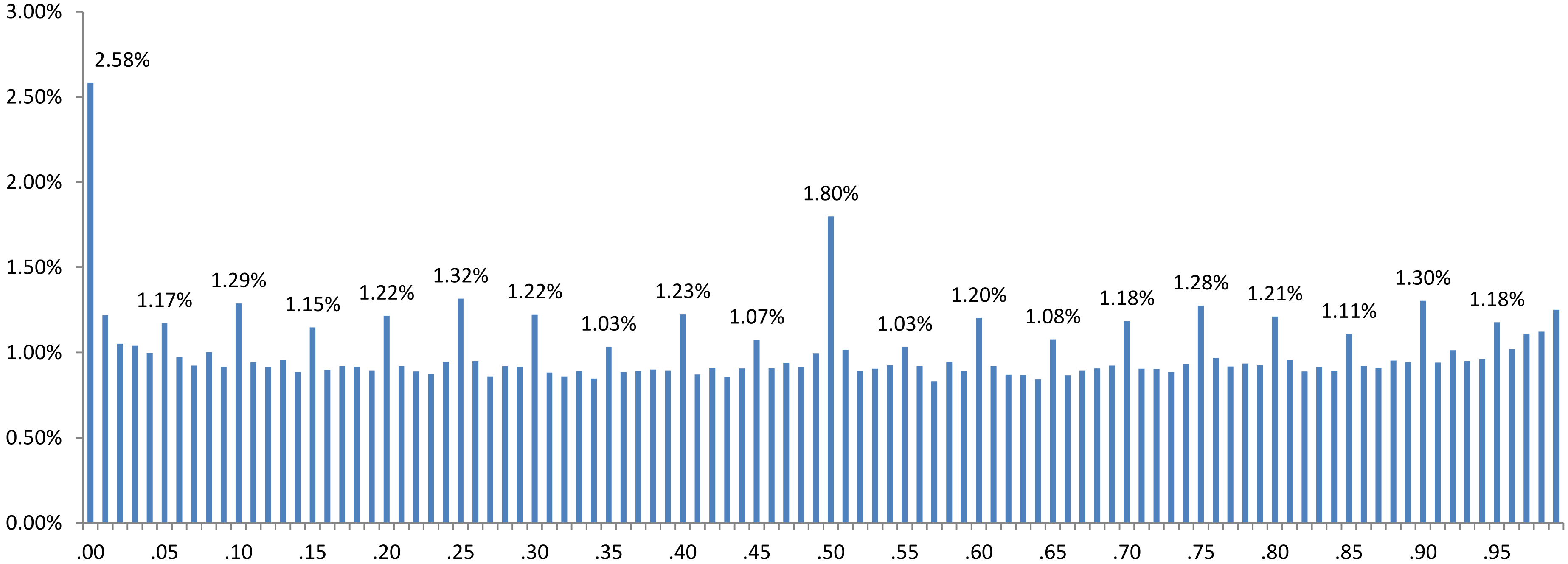


Figure 4: ‘Round-Number’ Bias: ‘Discount’ vs ‘Full-service’ brokers (AN sample, 2001-2011)

Following KLZ (2015), we first round the transaction price of all sub-sample trades to the nearest penny, e.g. ‘X.01’, ‘X.02’ and etc. Then we aggregate them by number at a given ‘rounded price’ (e.g. ‘X.01’) and scale by the total number of trades conducted by all sample institutions within a trading day to get the daily ‘round ratio’. And finally we calculate the arithmetic mean across all sample trading days (Feb.2001-Sep.2011) for a given type of brokers and plot accordingly. Here, we separate the Abel Noser sample trades into two groups based on the per share commission: those brokers whose in-sample per share commissions above (below) 3 cents are classified as being executed by ‘full-service’ (‘discount’) brokers, as in Anand et al. (2012). And we only include trades that can be matched to the ‘PERMNO’ code in the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3, and trades submitted by institutions with client type being 1 or 2, i.e. plan sponsors and investment managers. Fig. 4A reports the average daily ‘round ratio’ based on trades executed by ‘discount’ brokers and Fig. 4B shows the result based on the subsample of Abel Noser trades executed by ‘full-service’ brokers. The sample starts from Feb. 2001 to Sep. 2011.

Fig. 4A: 'Round-number bias' of 'discount' brokers (based on the number of trades)

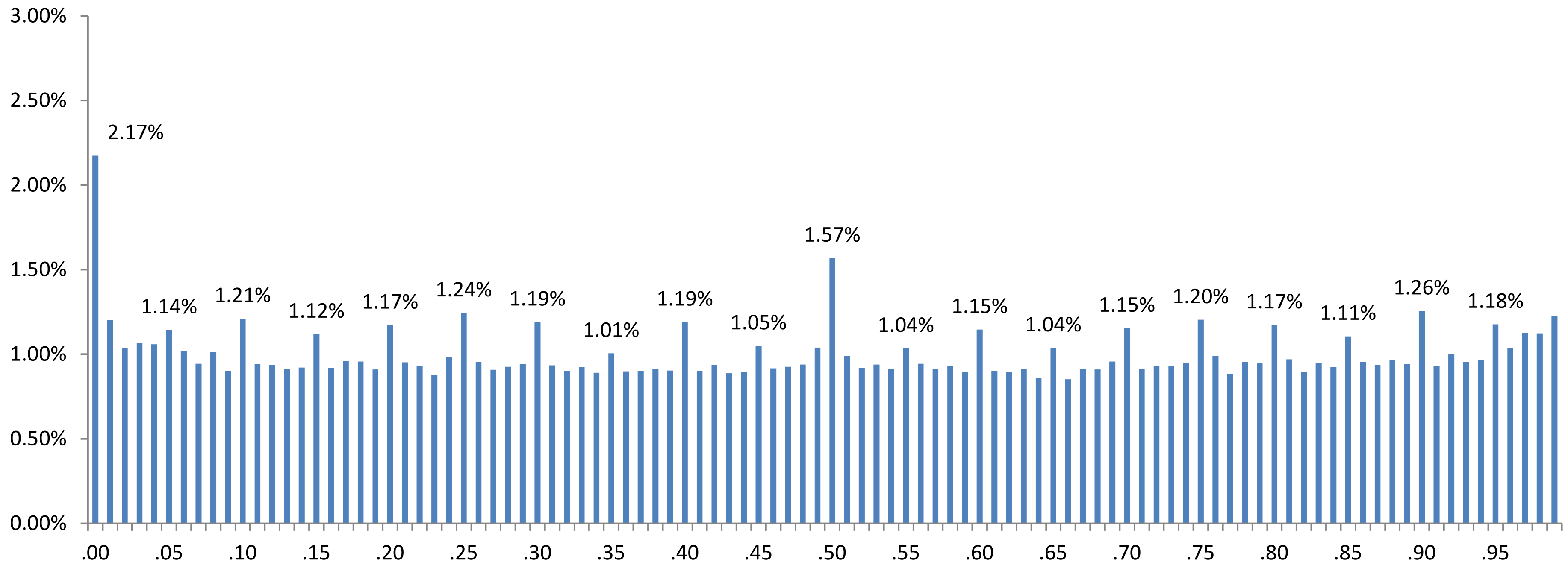


Fig. 4B: 'Round-number bias' of 'full-service' brokers (based on the number of trades)

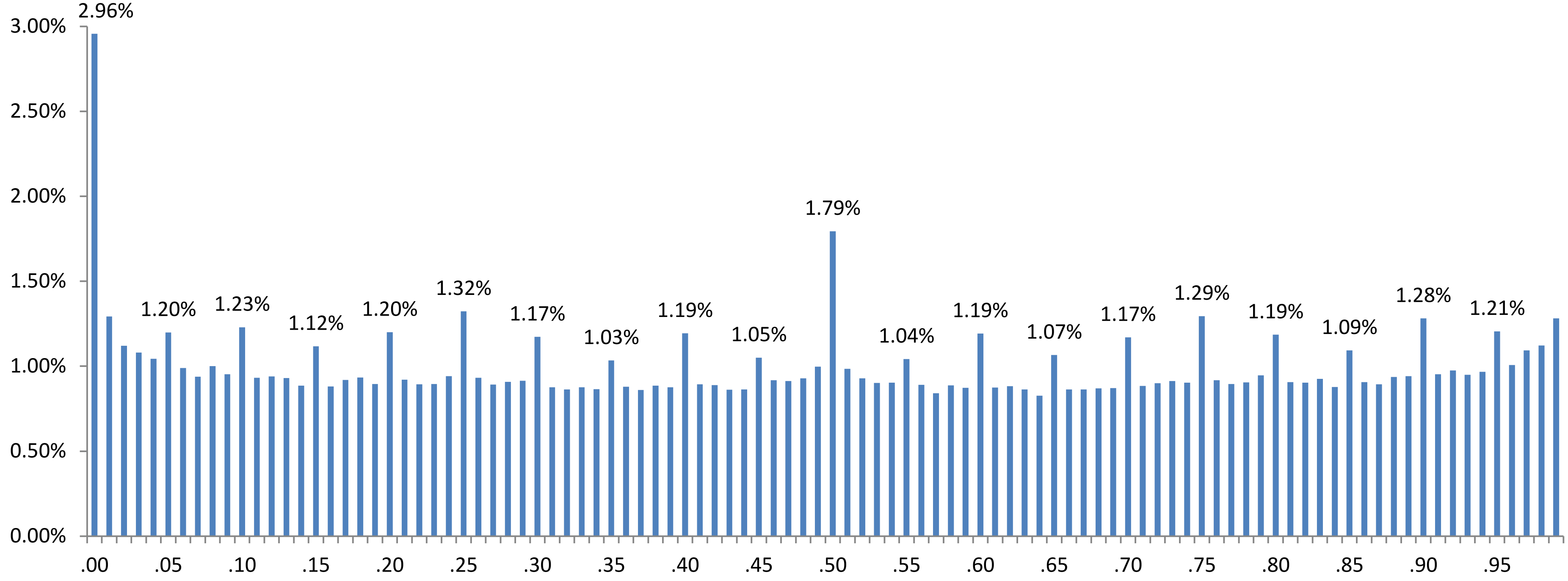


Figure 5: 'Round-Number' Bias: 'Large' vs 'Small' institutions (AN sample, 2001-2011)

Following KLZ (2015), we first round the transaction price of all sample institutions' trades to the nearest penny, e.g. 'X.01', 'X.02' and etc. Then we aggregate them by number for a given 'rounded price' (e.g. 'X.01') within the same type of institutions in a trading day to get the 'round ratio' for each rounded price. And finally we plot the arithmetic mean across all sample days (2001-2011) for a given type of institutions. Here, we separate the Abel Noser institutions (with unique 'clientcode') into two groups based on their total trading volume (in dollars) within the overall sample (2001-2011). Those above the top 33th percentile are classified as 'large' institutions and those below the bottom 33th percentile are labelled as 'small' institutions. And we only include trades that can be matched to the 'PERMNO' code from the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3 and submitted by institutions with client type being 1 or 2, i.e. plan sponsors and investment managers. Fig. 5A reports the average daily 'round ratio' within the subset of Abel Noser trades submitted by 'large' institutions and Fig. 5B shows the average daily 'round ratio' within trades submitted by 'small' institutions. The sample starts from Feb. 2001 to Sep. 2011.

Fig. 5A: 'Round-number bias' of 'large' institutions (based on the number of trades)

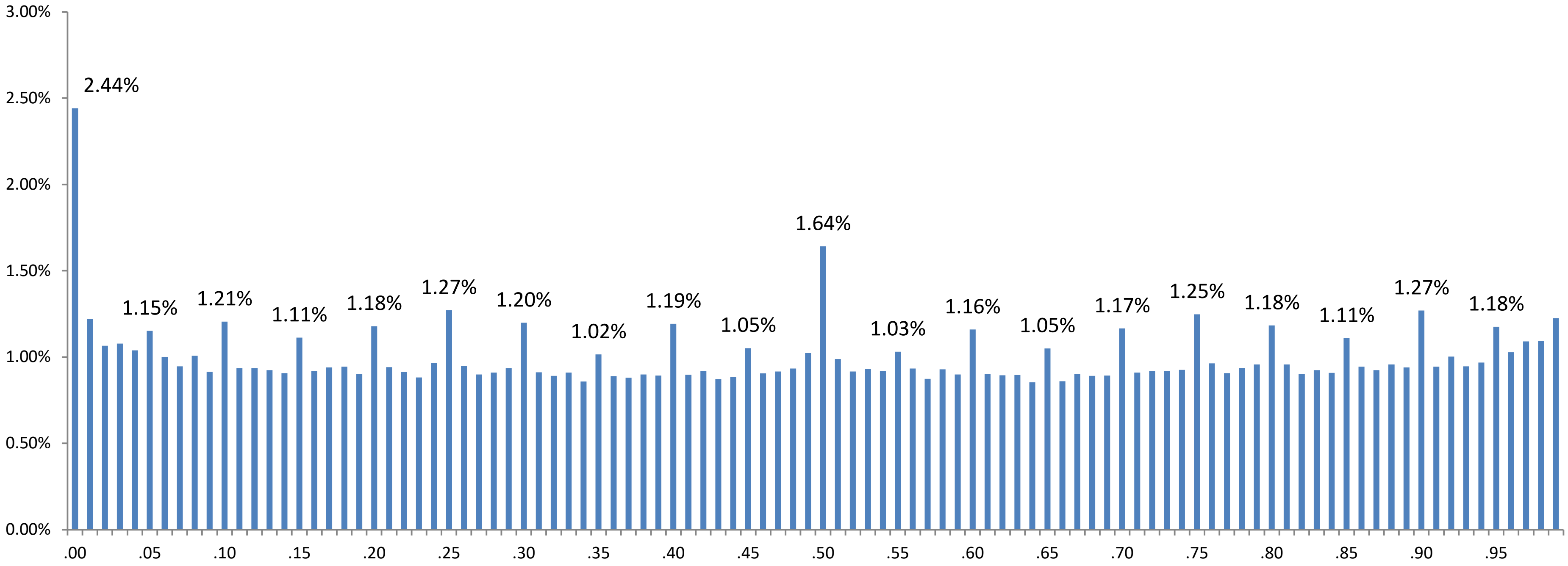


Fig. 5B: 'Round-number bias' of 'small' institutions (based on the number of trades)

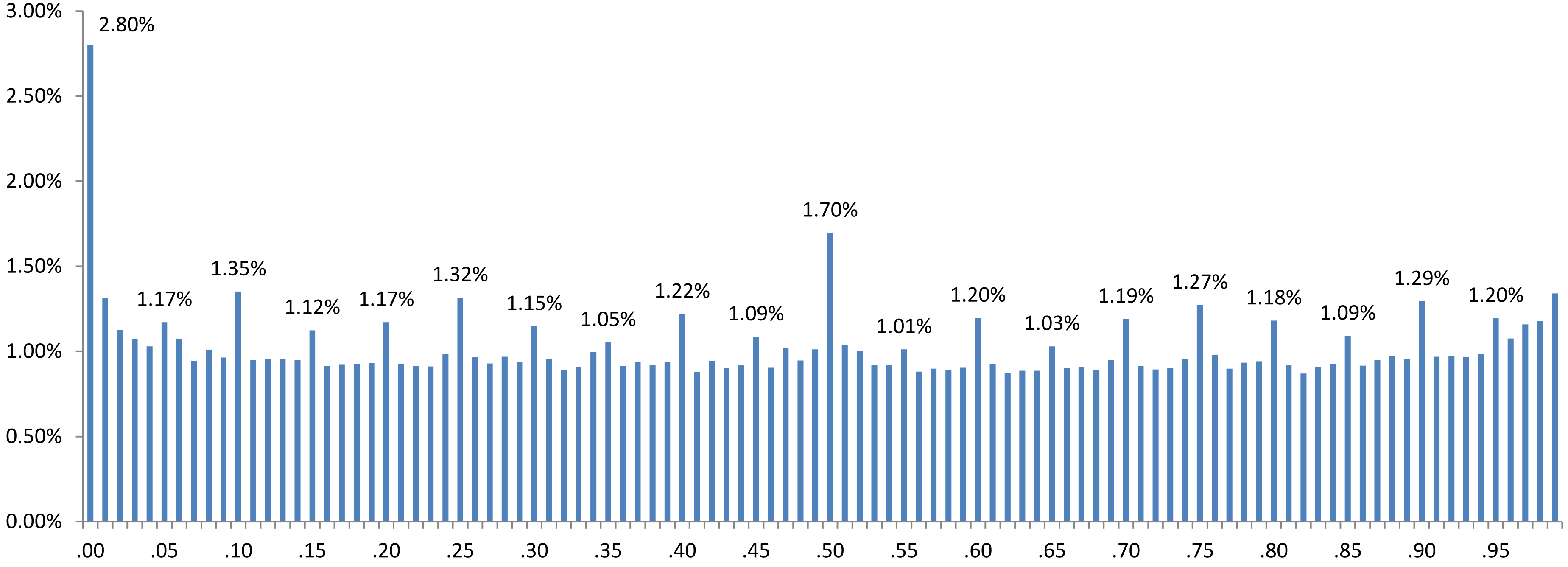


Table 4: Regression estimations of the difference of the 'round-number' bias across different samples: AN vs. TAQ sample (2001-2011)

Dependent Variable: 'Round Ratio' (%)	(1) TAQ sample	(2) AN sample	(3) AN vs TAQ
Intercept	0.89^{***}	1.01^{***}	0.95^{***}
p-value	<*.0001*	<*.0001*	<*.0001*
D_{00}	2.56^{***}	1.52^{***}	2.47^{***}
p-value	<*.0001*	<*.0001*	<*.0001*
D_{50}	1.15^{***}	0.70^{***}	1.12^{***}
p-value	<*.0001*	<*.0001*	<*.0001*
D_{X0}	0.57^{***}	0.25^{***}	0.54^{***}
p-value	<*.0001*	<*.0001*	<*.0001*
D_{X5}	0.42^{***}	0.20^{***}	0.40^{***}
p-value	<*.0001*	<*.0001*	<*.0001*
Indicator			0.05^{***}
p-value			<*.0001*
$D_{00}\times$ Indicator			-0.95^{***}
p-value			<*.0001*
$D_{50}\times$ Indicator			-0.42^{***}
p-value			<*.0001*
$D_{X0}\times$ Indicator			-0.29^{***}
p-value			<*.0001*
$D_{05}\times$ Indicator			-0.21^{***}
p-value			<*.0001*
Average R^2	3.49%	4.82%	3.75%
N	2651	2651	2651

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

p-value in Italic.

For each 'rounded penny' price, we compute the ratio between the number of trades executed at that price and the total number of trades executed in the same trading day within the AN sample (Column 1) or the TAQ sample (Column 2). Then, we regress these 'round ratios' (100 in total) on a series of 'roundness' indicators (e.g. D_{X0} equals 1 for the 'rounded penny' price '.X0' where X is an integer ranging from 1 to 9, excluding 5) and different interaction terms (Column 3), as in Equation (2) in KLZ (2015), using the Fama-MacBeth approach. In column (3) the interaction term (Indicator) equals 1 if the 'round ratio' is from the Abel Noser sample. And we only include trades that can be matched to the 'PERMNO' code from the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3 and submitted by institutions with client type being 1 or 2, i.e. plan sponsors and investment managers. The Abel Noser sample starts from Apr. 2001 to Sep. 2011 and the TAQ sample starts from Apr. 2001 and ends in Dec. 2014.

Table 5: Regression estimations of the difference of the 'round-number' bias across subsamples in Abel Noser (2001-2011)

Dependent Variable: 'Round Ratio' (%)	(1) 'Trader' vs Non-'trader'	(2) 'Discount' vs 'Full-service'	(3) 'Big' vs 'Small'
Intercept	0.93***	1.03***	0.93***
p-value	<*.0001*	<*.0001*	<*.0001*
D_{00}	2.04***	2.02***	1.72***
p-value	<*.0001*	<*.0001*	<*.0001*
D_{50}	0.88***	0.86***	0.75***
p-value	<*.0001*	<*.0001*	<*.0001*
D_{X0}	0.26***	0.27***	0. 24***
p-value	<*.0001*	<*.0001*	<*.0001*
D_{X5}	0.22***	0.22***	0.17***
p-value	<*.0001*	<*.0001*	<*.0001*
Indicator	-0.01***	0.01**	0. 00
p-value	<*.0001*	*0.018*	*0.4505*
$D_{00}\times$ Indicator	-0.81***	-0.80***	-0.26***
p-value	<*.0001*	<*.0001*	<*.0001*
$D_{50}\times$ Indicator	-0.74***	-0.28***	-0.03
p-value	<*.0001*	<*.0001*	*0.3918*
$D_{X0}\times$ Indicator	-0.96***	-0.04***	0.01
p-value	<*.0001*	*0.0002*	*0.4037*
$D_{05}\times$ Indicator	-0.96***	-0.05***	0.01
p-value	<*.0001*	<*.0001*	*0.2821*
Average R^2	35%	80%	13%
N	2651	2651	2651

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

p-value in Italic.

For each 'rounded penny' price, we compute the ratio between the number of trades executed at that price and the total number of trades executed in the same trading day within the same sub-group (e.g. the total number of trades submitted by a 'discount' broker or those that can be matched with 'valid' trader codes in the AN sample). Then, we regress these 'round ratios' (200 in total) on a series of 'roundness' indicators (e.g. D_{X0} equals 1 for the 'rounded penny' price '.X0' where X is an integer ranging from 1 to 9, excluding 5) and different interaction terms, as in Equation (2) in KLZ (2015), using the Fama-MacBeth approach. In column (1) the interaction term (Indicator) equals 1 if the transaction is from the Abel Noser sample. In column (4) the Indicator denotes a trade from the Abel Noser that has a 'valid' trader code. And the Indicator in column (5) equals 1 if the per share commission of a transaction from the Abel Noser dataset is no more than 3 cents, according to Anand et al. (2012). In column (6), the Indicator equals 1 if the total dollar trading volume of a sample institution is above the 66th percentile in our AN sample (and equals 0 if it is below the 33th percentile), aggregated from Feb.2001 to Sep. 2011. The TAQ sample starts from Apr. 2001 to Sep. 2011 and the TAQ sample starts from Apr. 2001 and ends in Dec. 2014.

Table 6: Regression estimates for the annual change of the 'round number' bias in the AN sample (2001-2011) and the TAQ sample (2001-2014)

Dependent Variables	TAQ (1) 'Round Ratio' (%) (N)	TAQ (2) 'Round Ratio' (%) (D)	AN (3) 'Round Ratio' (%) (N)	AN (4) 'Round Ratio' (%) (D)
D_00	2.78***	4.80***	3.42***	2.86***
D_50	1.98***	3.15***	2.11***	1.87***
D_X0	1.53***	2.06***	1.37***	1.25***
D_X5	1.35***	1.69***	1.26***	1.68***
Year	0.01***	0.02***	0.01***	0.01***
D_00×Year	-0.09***	-0.23***	-0.20***	-0.14***
D_50×Year	-0.05***	-0.14***	-0.09***	-0.07***
D_X0×Year	-0.04***	-0.08***	-0.04***	-0.02***
D_X5×Year	-0.03***	-0.05***	-0.03***	-0.01***
Intercept	0.87***	0.75***	0.89***	0.92***
Day-of-week fixed effect	Yes	Yes	Yes	Yes
Adj. R^2	0.9875	0.9744	0.5949	0.6873
N	345900	345900	265100	265100

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

For each sample trade, we first round the actual trade price (*p*) to the nearest penny level. The daily 'round ratio' equals the number of trades/shares/dollars traded at a given 'rounded price' (e.g. 'X.01') divided by the total number/shares/dollars of trades within a trading day. The dependent variables in column (1)&(3) and (2)&(4) are the 'round ratio' at different 'rounded penny' prices (100 in total) aggregated by the number of trades or dollars traded, respectively. Then we regress the daily VW 'round ratio' of a given 'rounded penny' price on a series of dummy variables, i.e. D_{00}, D_{50}, D_{X0} and D_{X5}, with six dummy variables controlling for the day of the week. We estimate the annual time trend by multiplying the 3 dummy variables for different types of 'rounded penny' prices with the time length term (number of years since the beginning of the sample). All estimated coefficients and intercepts are reported in percentage (1%). The AN sample starts from Apr. 2001 and ends in Sep. 2011 and the TAQ sample ends in Dec. 2014.

Table 7: Tests for the difference of the average 'round ratio' before and after the 'auto-quote' event in NYSE (2003)

Panel A: Univariate Analysis (*t*-test)							
Event:	**The adoption of 'auto-quote' program in NYSE (Jan. 29 to May 27, 2003)**						
'Round Ratio' (%)	NYSE			NASDAQ			Diff-in-Diff
	Pre-event	Post-event	Post-Pre (T)	Pre-event	Post-event	Post-Pre (C)	(T-C)
'.00'	2.44	2.29	-0.16***	2.59	2.63	0.04	-0.20**
'.50'	1.83	1.73	-0.10***	1.80	2.00	0.19***	-0.29***
'.X0'	1.50	1.41	-0.08***	1.48	1.52	0.04***	-0.12***
'.X5'	1.32	1.26	-0.06***	1.33	1.34	0.01	-0.07***

Panel B: Regression Analysis*

Dependent variable: 'Round Ratio' (%)

	'.00'	**'.50'**	**'.X0'**	**'.X5'**
$Treatment$	-0.2335**	-0.0919	0.0478*	0.0057
$Post$	-0.0263	0.1264***	0.0554***	0.0184
$\boldsymbol{Treatment * Post}$	**-0.1780***	**-0.2454*****	**-0.1546*****	**-0.0856*****
$\log(volume)$	0.1573	0.0777	0.1315***	0.0512
VW_RET	1.8623	1.1574	-0.0034	-0.0628
VW_Range	0.0007	0.0057	0.0041	0.0054
VW_Amihud	-0.0992	-1.3568	2.0017**	0.9225
Day-of-week fixed effect	Yes	Yes	Yes	Yes
N (Trading days)	165	165	165	165
Adj. R-square	*24.06*%	*27.29*%	*7.19*%	*3.55*%

* Regression model: $y_{i,t} = \beta_1 \cdot Treatment + \beta_2 \cdot Post + \beta_3 \cdot Treatment * Post - event + \beta_4 \cdot \log(volume)_t + \beta_5 \cdot VW\_RET_t + \beta_6 \cdot VW\_Range_t + \beta_7 \cdot VW\_Amihud_t + e_{i,t}$

$^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Following Hendershott, Jones and Menkveld (2011), here we construct the daily 'panel' of our sample of NYSE and NASDAQ TAQ trades using data two months before and after the event period. For each sample transaction of the TAQ database, we first round the transaction price to the nearest penny, e.g. 'X.00', 'X.50'. Then we aggregate them and count the number of trades at a given 'rounded price' (e.g. 'X.00') and scale by the total number of transactions within a trading day. Finally, we compute the daily 'round ratio' for each 'rounded penny' price by computing the VW mean of the 'round ratio' using the total number of trades executed for each sample stock. log*(volume)* is the natural log of the total daily trading volume (in dollars) in either NYSE or NASDAQ. 'VW_RET' refers to the weighted average returns of stocks traded at a stock exchange during a sample trading day, based on the market capitalization at the close time. 'VW_Amihud' is the daily volume-weighted average of the Amihud's 'illiquidity' measure within a stock exchange, which equals the ratio between the absolute daily return and the trading volume then multiplied by 1×10^6. And 'VW_Range' is the volume-weighted average of the daily price range within the same exchange, as used in Gai, Yao and Ye (2014). And we include a series of dummy variables indicating the day of week. Here, we only include trades that can be matched to the 'PERMNO' code from the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1 or 3 within the normal trading hour, i.e. from 9:30 a.m. to 4:00 p.m. in Eastern Time.

Table 8: Tests for the change in the average 'effective spread' at 'round-number' prices before and after the 'auto-quote' event in NYSE (2003)

Panel A: Univariate Analysis (*t*-test)							
Event:	**The adoption of 'auto-quote' program in NYSE (Jan. 29 to May 27, 2003)**						
'Effective Cost' (in bps)	NYSE Pre-event	NYSE Post-event	NYSE Post-Pre (T)	NASDAQ Pre-event	NASDAQ Post-event	NASDAQ Post-Pre (C)	Diff-in-Diff (T-C)
'.00'	6.0	4.9	-1.1***	7.9	7.8	-0.1	-1.0*
'.50'	5.8	4.4	-1.4***	7.7	7.6	0.1	-1.5**
'.X0'	5.4	4.1	-1.3***	6.7	6.6	-0.1	-1.2***
'.X5'	5.1	3.8	-1.4***	6.4	6.4	0.0	-1.4***

Panel B: Regression Analysis*

Dependent variable: VW 'effective cost' (in bps)

	'.00'	**'.50'**	**'.X0'**	**'.X5'**
$Treatment$	-2.7377***	-2.1308***	-1.8747***	-1.6748***
$Post$	-0.6771	-0.4245	-0.3813***	-0.2041
$\boldsymbol{Treatment * Post}$	**-1.0545**	**-1.4284****	**-1.2850*****	**-1.2878*****
$\log(volume)$	3.2032***	2.3933**	2.1296***	0.7767
VW_RET	-8.6151	6.1111	-6.2253**	-2.8223
VW_Range	0.0124	-0.0394	0.1137	0.0850
VW_Amihud	17.4905	21.4271	13.5368	3.1156
Day-of-week fixed effect	Yes	Yes	Yes	Yes
N (Trading days)	165	165	165	165
Adj. R-square	*40.04*%	*41.87*%	*51.86*%	*15.83*%

***** Regression model: $y_{i,t} = \beta_1 \cdot Treatment + \beta_2 \cdot Post + \beta_3 \cdot Treatment * Post + \beta_4 \cdot \log(volume)_t + \beta_5 \cdot VW\_RET_t + \beta_6 \cdot VW\_Range_t + \beta_7 \cdot VW\_Amihud_t + e_{i,t}$

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Following Hendershott, Jones and Menkveld (2011), here we construct the daily panel of our sample of NYSE and NASDAQ TAQ trades using data two months before and after the event period. For each sample trade, we first round the actual trade price (p) to the nearest penny level. Then we match the mid-point of the NBBO quote (m) provided by the WRDS database at the same time (with 0-second delay). Following Hasbrouck (2009), the 'effective cost' measure is defined as the difference between the log transaction price and the log mid-point quote, multiplied by (-1) if it is a sell trade. We apply the Lee & Ready (2001) algorithm to determine the direction of each sample trade. And finally we compute the daily VW 'effective cost' measure based on all sample trades executed at either '.00' or '.50', using the dollar trading volume as the weight. Here, we only include trades that can be matched to the 'PERMNO' code from the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1 or 3 within the normal trading hour, i.e. from 9:30 a.m. to 4:00 p.m. in Eastern Time. *log(volume)* is the natural log of the total daily trading volume (in dollars) in NYSE or NASDAQ exchange. 'VW_RET' refers to the weighted average returns of stocks traded at either NYSE or NASDAQ during a sample trading day, based on the market capitalization at the close time. 'VW_Amihud' is the daily volume-weighted average of the Amihud's 'illiquidity' measure, which equals the ratio between the absolute daily return and the trading volume then multiplied by 1×10^6. And 'VW_Range' is the volume-weighted average of the daily price range within the same exchange, as used in Gai, Yao and Ye (2014).

Table 9: Regression estimates for the difference of the 'effective cost' of different types of 'rounded penny' prices (the TAQ sample, 2001-2014)

Dependent Variable: Effective Cost (in bps)	(1) VW	(2) EW
Intercept	4.6^{***}	13.0^{***}
p-value	*<.0001*	*<.0001*
D_{00}	1.7^{***}	8.0^{***}
p-value	*<.0001*	*<.0001*
D_{50}	1.8^{***}	7.7^{***}
p-value	*<.0001*	*<.0001*
D_{X0}	1.0^{***}	4.0^{***}
p-value	*<.0001*	*<.0001*
D_{X5}	0.9^{***}	3.2^{***}
p-value	*<.0001*	*<.0001*
Average R^2	1.11%	5.05%
F-tests:		
D_{00}-D_{X0}	*116.19*	*1032.55*
D_{50}-D_{X0}	*146.14*	*872.72*
D_{00}-D_{X5}	*139.34*	*1429.72*
D_{50}-D_{X5}	*172.26*	*1237.88*
D_{00}-D_{50}	*0.96*	*3.78*
N	345900	345900

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

p-value in Italic.

For each sample TAQ trade, we first round the actual trade price (*p*) to the nearest penny level. Then we match the mid-point of the NBBO quote (m) provided by the WRDS database at the same time (with 0-second delay). Following Hasbrouck (2009), the 'effective cost' measure is defined as the difference between the log transaction price and the log mid-point quote, multiplied by (-1) if it is a sell trade. We apply the Lee & Ready (2001) algorithm to determine the direction of each sample trade. And finally we compute the daily VW 'effective cost' measure (in bps, column 1) based on all sample trades executed at a given 'rounded penny' price, using the dollar trading volume as the weight. We also compute the daily EW average 'effective cost' across all sample stocks at a given 'rounded penny' price (in bps, column 2). Finally, we regress the daily VW/EW 'effective cost' measure on a series of dummy variables, i.e. $D_{00}/D_{50}/D_{X0}/D_{X5}$, using the Fama-MacBeth approach. In the 3 rows below the 'Average R^2', we conduct the *F*-test by comparing the estimated coefficients among D_{00}, D_{50} and D_{X0}. Here, we only include trades that can be matched to the 'PERMNO' code in the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1 or 3 within the normal trading hour, i.e. from 9:30 a.m. to 4:00 p.m. in Eastern Time. The sample period starts from Feb. 2001 to Dec.2014.

Table 10: Estimations of the economic impact ('effective cost') of the sample TAQ trade executed at different types of rounded prices (2001-2014)

Panel A:	**Average 'effective cost' (in bps): 'round-number' vs non 'round-number' prices**													
Year	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**Full-sample**
'.00'	14.82	11.11	7.82	6.29	5.19	4.70	4.90	6.24	5.50	4.00	4.19	4.17	5.15	4.96	6.35
'.50'	15.43	11.56	8.04	6.49	5.18	4.84	4.67	6.98	5.53	4.05	4.22	4.12	4.59	4.70	6.44
'.X0'	12.04	9.82	6.94	5.52	4.51	4.25	4.22	5.65	5.15	3.80	3.99	3.88	4.64	4.28	5.70
'.X5'	11.99	9.59	6.77	5.35	4.39	4.12	4.13	5.54	5.22	3.73	3.93	3.92	4.90	4.18	5.53
Any other price	8.98	7.13	5.15	4.27	3.66	3.50	3.54	4.90	4.69	3.43	3.69	3.62	4.45	3.90	4.63
Panel B:	**Annual dollar volume (in billions): 'round-number' vs non 'round-number' prices**														
Year	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**Full-sample**
'.00'	1,321	755	541	603	678	750	1,001	1,094	614	677	758	743	731	878	11,144
'.50'	676	492	368	407	456	502	654	707	444	488	534	486	502	603	7,319
'.X0'	313	312	262	290	323	352	455	479	343	375	392	352	368	430	5,047
'.X5'	291	246	217	244	273	303	400	431	313	342	357	318	333	391	4,457
Any other price	108	98	109	135	162	198	286	332	252	277	291	259	267	317	3,091
Panel C:	**Additional 'Effective Cost' (in millions) caused by the 'round-number' bias**														
Year	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**Full-sample**
'.00'	772	301	144	122	104	90	136	147	50	39	38	41	51	93	1,918
'.50'	436	218	106	90	70	67	74	147	38	30	28	25	7	48	1,326
'.X0'	96	84	47	36	28	27	31	36	16	14	12	9	7	16	488
'.X5'	88	61	35	26	20	19	23	28	17	10	9	10	15	11	399
Total amount	2,854	1,798	974	764	594	558	692	861	381	286	247	237	262	376	11,138

* Here, to estimate the additional cost of certain 'round-number' prices, we first round the actual trade price (*p*) to the nearest penny level for each sample TAQ trade. Then we match the mid-point of the NBBO quote (*m*) provided by the WRDS database at the same time (with 0-second delay). Following Hasbrouck (2009), the 'effective cost' measure is defined as the difference between the log transaction price and the log mid-point quote, multiplied by 1 (-1) if it is a buy (sell) trade. We apply the Lee & Ready (2001) algorithm to determine the direction of each sample trade. Next we compute the daily VW 'effective cost' measure (%) based on all sample trades executed at a given 'rounded penny' price within the same trading day, using the dollar trading volume as the weight. In Panel A, we report the estimated daily 'effective cost' (%) for each type of 'rounded penny' price by running the Fama-MacBeth type regression with a series of dummy variables: D_{00}, D_{50}, D_{X0} and D_{X5} and the intercept (include prices other than '.00', '.50', '.X0' and '.X5', termed as 'Any other price') within each calendar year. In Panel B, we report the total dollar volume of the sample TAQ trade executed at each type of 'rounded penny' price aggregated at each sample year. And in Panel C, we report the additional 'effective cost' (in millions) for those sample TAQ trades executed at a given type of 'rounded penny' price by multiplying the gap between the VW 'effective cost' (%) of the trade executed at a given 'rounded penny' price and that at 'Any other price' with the annual total dollar trading volume of the trade executed at the given 'rounded penny' price. Here, 'Any other price' refers to the average 'effective cost' (or the total dollar trading volume in Panel B) of the sample TAQ trade executed at a price other than '.00', '.50', '.X0' or '.X5'. The TAQ sample period ranges from Feb. 2001 to Dec. 2014.

Reference List

Aitken, M., Brown, P., Buckland, C., Izan, H. Y., & Walter, T. (1996). Price clustering on the Australian stock exchange. *Pacific-Basin Finance Journal, 4*(2-3), 297-314.

Ahn, H. J., Cao, C. Q., & Choe, H. (1996). Tick size, spread, and volume. *Journal of Financial Intermediation*, *5*(1), 2-22.

Alexander, G. J., & Peterson, M. A. (2007). An analysis of trade-size clustering and its relation to stealth trading. *Journal of Financial Economics*, *84*(2), 435-471.

Anand, A., Irvine, P., Puckett, A., & Venkataraman, K. (2012). Performance of institutional trading desks: An analysis of persistence in trading costs. *The Review of Financial Studies*, *25*(2), 557-598.

Bai, J., Philippon, T., & Savov, A. (2016). Have financial markets become more informative?. *Journal of Financial Economics*, *122*(3), 625-654.

Ball, C. A., Torous, W. N., & Tschoegl, A. E. (1985). The degree of price resolution: The case of the gold market. *Journal of Futures Markets*, *5*(1), 29-43.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2007). Is the aggregate investor reluctant to realise losses? Evidence from Taiwan. *European Financial Management*, *13*(3), 423-447.

Barber, B. M., Odean, T., & Zhu, N. (2009). Systematic noise. *Journal of Financial Markets*, *12*(4), 547-569.

Baker, M., Litov, L., Wachter, J. A., & Wurgler, J. (2010). Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements. *Journal of Financial and Quantitative Analysis*, *45*(5), 1111-1131.

Bessembinder, H. (1999). Trade execution costs on Nasdaq and the NYSE: A post-reform comparison. *Journal of Financial and Quantitative Analysis*, *34*(3), 387-407.

Bessembinder, H. (2003). Trade execution costs and market quality after decimalization. *Journal of Financial and Quantitative Analysis*, *38*(4), 747-777.

Bhattacharya, U., Holden, C. W., & Jacobsen, S. (2012). Penny wise, dollar foolish: Buy–sell imbalances on and around round numbers. *Management Science*, *58*(2), 413-431.

Bhattacharya, U., Kuo, W. Y., Lin, T. C., & Zhao, J. (2018). Do superstitious traders lose money?. *Management Science*, *64*(8), 3772-3791.

Blume, M. E. (1993). Soft dollars and the brokerage industry. *Financial Analysts Journal*, *49*(2), 36-44.

Boehmer, E., Jones, C. M., Zhang, X., & Zhang, X. (2019). Tracking retail investor activity. *Working Paper*.

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *The Review of Financial Studies*, *27*(8), 2267-2306.

Brown, P., Chua, A., & Mitchell, J. (2002). The influence of cultural factors on price clustering: Evidence from Asia–Pacific stock markets. *Pacific-Basin Finance Journal*, *10*(3), 307-332.

Brown, P., & Mitchell, J. (2008). Culture and stock price clustering: Evidence from The Peoples'

Republic of China. *Pacific-Basin Finance Journal*, *16*(1-2), 95-120.

Busse, J. A., Chordia, T., Jiang, L., & Tang, Y. (2019). Transaction Costs, Portfolio Characteristics, and Mutual Fund Performance. *Management Science*, Forthcoming.

Buti, S., Rindi, B., Wen, Y., & Werner, I. M. (2013). Tick size regulation and sub-penny trading. *Working Paper*, *Charles A. Dice Center*, The Ohio State University.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.

Chemmanur, T. J., He, S., & Hu, G. (2009). The role of institutional investors in seasoned equity offerings. *Journal of Financial Economics*, *94*(3), 384-411.

Chen, H. L., Jegadeesh, N., & Wermers, R. (2000). The value of active mutual fund management: An examination of the stockholdings and trades of fund managers. *Journal of Financial and Quantitative Analysis*, *35*(3), 343-368.

Chiang, Y.M., Hirshleifer, D., Qian Y., & Sherman, A. (2011). Do investors learn from experience? Evidence from frequent IPO investors. *Review of Financial Studies*, *24(5)*, 1560–1589.

Chung, K. H., Van Ness, B. F., & Van Ness, R. A. (2004). Trading costs and quote clustering on the NYSE and NASDAQ after decimalization. *Journal of Financial Research*, *27*(3), 309-328.

Chakrabarty, B., & Chung, K. H. (2004). Can sub-penny pricing reduce trading costs?. Working Paper, Saint Louis University.

Chordia, T., Roll, R., & Subrahmanyam, A. (2011). Recent trends in trading activity and market quality. *Journal of Financial Economics*, *101*(2), 243-263.

Christie, W. G., Harris, J. H., & Schultz, P. H. (1994). Why did NASDAQ market makers stop avoiding odd-eighth quotes?. *The Journal of Finance*, *49*(5), 1841-1860.

Chung, H. M., & Chiang, S. M. (2006). Price clustering in E-mini and floor-traded index futures. *Jouranl of Futures Markets*, 26 (3), 269-295.

Coulter, K. S., & Coulter, R. A. (2007). Distortion of price discount perceptions: The right digit effect. *Journal of Consumer Research*, *34*(2), 162-173.

Coval, J. D., & Shumway, T. (2005). Do behavioral biases affect prices?. *The Journal of Finance*, *60*(1), 1-34.

Curcio, R., & Goodhart, C. (1991). The clustering of bid/ask prices and the spread in the foreign exchange market. *Working Paper*, London School of Economics.

Edelen, R. M., Ince, O. S., & Kadlec, G. B. (2016). Institutional investors and stock return anomalies. *Journal of Financial Economics*, *119*(3), 472-488.

Fama, E. F., & French, K. (2008). Mutual fund performance. *The Journal of Finance*, *63*(1), 389-416.

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, *61*(4), 2017-2046.

Frazzini, A., & Lamont, O. A. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics*, *88*(2), 299-322.

Fritz, F. (2014). Price clustering and round number effects-evidence from the German market. *Working Paper, 2014 Financial Markets & Corporate Governance Conference*.

Foucault, T. (2012). Algorithmic trading: Issues and preliminary evidence. *Market Microstructure (Part I Economic Microstructure Theory)*. John Wiley & Sons.

Foucault, T., Kadan, O., & Kandel, E. (2005). Limit order book as a market for liquidity. *The Review of Financial Studies*, *18*(4), 1171-1217.

Jain, P. K. (2005). Financial market design and the equity premium: Electronic versus floor trading. *The Journal of Finance*, *60*(6), 2955-2985.

Jones, C. M., & Lipson, M. L. (2001). Sixteenths: direct evidence on institutional execution costs. *Journal of Financial Economics*, *59*(2), 253-278.

Goetzmann, W. N., Kim, D., Kumar, A., & Wang, Q. (2015). Weather-induced mood, institutional investors, and stock returns. *The Review of Financial Studies*, *28*(1), 73-111.

Goldstein, M. A., Irvine, P., Kandel, E., & Wiener, Z. (2009). Brokerage commissions and institutional trading patterns. *The Review of Financial Studies*, *22*(12), 5175-5212.

Goldstein, M. A., & Kavajecz, K. A. (2000). Eighths, sixteenths, and market depth: changes in tick size and liquidity provision on the NYSE. *Journal of Financial Economics*, 56(1), 125-149.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2012). IQ, trading behavior, and performance. *Journal of Financial Economics*, *104*(2), 339-362.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, *70*(3), 393-408.

Hagströmer, B., & Nordén, L. (2013). The diversity of high-frequency traders. *Journal of Financial Markets*, 16(4), 741-770.

Hasbrouck, J. (2009). Trading costs and returns for US equities: Estimating effective costs from daily data. *The Journal of Finance*, 64(3), 1445-1477.

Hasbrouck, J., & Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, 16(4), 646-679.

Harris, L. (1991). Stock price clustering and discreteness. *The Review of Financial Studies*, 4(3), 389-415.

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity?. *The Journal of Finance*, 66(1), 1-33.

Hendershott, T., & Riordan, R. (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis*, 48(4), 1001-1024.

Hu, G. (2009). Measures of implicit trading costs and buy–sell asymmetry. *Journal of Financial Markets*, 12(3), 418-437.

Hu, G., Jo, K. M., Wang, Y. A., & Xie, J. (2018). Institutional trading and Abel Noser data. *Journal of Corporate Finance*, 52, 143-167.

Hu, G., McLean, R. D., Pontiff, J., & Wang, Q. (2014). The year-end trading activities of institutional investors: Evidence from daily trades. *The Review of Financial Studies*, 27(5), 1593-1614.

Huang, R. D., & Stoll, H. R. (1996). Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. *Journal of Financial Economics*, *41*(3), 313-357.

Ikenberry, D. L., & Weston, J. P. (2008). Clustering in US stock prices after decimalisation. *European Financial Management*, 14(1), 30-54.

Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263-291.

Kacperczyk, M., Nieuwerburgh, S. V., & Veldkamp, L. (2014). Time-varying fund manager skill. *The Journal of Finance*, *69*(4), 1455-1484.

Kuo, W. Y., Lin, T. C., & Zhao, J. (2015). Cognitive limitation and investment performance: Evidence from limit order clustering. *The Review of Financial Studies*, 28(3), 838-875.

Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, 102(5), 2206-36.

Liu, H. C. (2011). Timing of price clustering and trader behavior in the foreign exchange market: evidence from Taiwan. *Journal of Economics and Finance*, *35*(2), 198-210.

Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4), 712-740.

Mitchell, J. (2001). Clustering and psychological barriers: The importance of numbers. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 21(5), 395-428.

Mitchell, J., & Izan, H. Y. (2006). Clustering and psychological barriers in exchange rates. *Journal of International Financial Markets, Institutions and Money*, *16*(4), 318-344.

Niederhoffer, V., & Osborne, M. F. M. (1966). Market making and reversal on the stock exchange. *Journal of the American Statistical Association*, 61(316), 897-916.

Puckett, A., & Yan, X. (2011). The interim trading skills of institutional investors. *The Journal of Finance*, 66(2), 601-633.

Repetto, L., & Solís, A. (2017). The price of inattention: Evidence from the Swedish housing market. *Journal of the European Economic Association*.

DeVault, L., Sias, R., & Starks, L. (2019). Sentiment metrics and investor demand. *The Journal of Finance*, *74*(2), 985-1024.

Seru, A., Shumway, T., & Stoffman, N. (2010). Learning by trading. *The Review of Financial Studies*, 23(2), 705-739.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99-118.

Sopranzetti, B. J., & Datar, V. (2002). Price clustering in foreign exchange spot markets. *Journal of Financial Markets*, *5*(4), 411-417.

Schindler, R. M., & Kirby, P. N. (1997). Patterns of rightmost digits used in advertised prices: implications for nine-ending effects. *Journal of Consumer Research*, *24*(2), 192-201.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, *4*(3), 199-214.

Thomas, M., & Morwitz, V. (2005). Penny wise and pound foolish: the left-digit effect in price cognition. *Journal of Consumer Research*, *32*(1), 54-64.

Xu, W. (2015). Active Institutional Investors and Stock Return Anomalies. Working Paper, Rutgers University.

Zhao, X., & Chung, K. H. (2006). Decimal pricing and information-based trading: tick size and informational efficiency of asset price. *Journal of Business Finance & Accounting*, *33*(5-6), 753-766.

Appendix

A1. The 'Left-digit' Effect among Institutional Traders in the AN Sample

In this section, we intend to investigate whether our sample institutions in the Abel Noser data also exhibit the 'left-digit' bias, i.e. they perceive those prices that are one penny below a certain 'round number' price (e.g. '.00', '.50') to be cheaper than the other type of prices and hence are more likely to buy the stock at such prices (which we refer to as 9-digit ending prices). This has been confirmed by Bhattacharya et al. (2012) using a cap-stratified sample of TAQ trades from 2001 to 2006. An advantage of the AN data comparing with the TAQ data is the label of the trade direction: as the former only (mostly) records one-side of the transaction, we do not need to rely on algorithms like the Lee-Ready or Huang and Stoll (1997). Following their methodology, in order to reduce the impact of 'outliers' with either extremely high/low 'round ratio (N)' for certain stocks or time periods, here I use the sample median accordingly. Here, in order to compare whether there is a heterogeneity across institutions in different size groups in terms of their 'left-digit' bias (if any), we aggregate the total number of buy and sell trades each calendar year at institution-level, instead of at stock-level as in Bhattacharya et al. (2012).

As shown in Panel A of Figure A1, we can see a very clear 'saw-shape' pattern of the median 'Buy-Sell Ratio' across all sample institutions along the 100 'rounded penny' prices: the median annual 'Buy-Sell Ratio' of an institution at those 9-digit ending prices is around 1.13, which is well above the unit level. And conversely, the figure also indicates that for a representative institution in our sample, those 1-digit ending prices are deemed as 'higher-selling' prices and thus receive more sell orders than the buy order (the 'Buy-Sell Ratio' is well below 1, 0.95). And in Panel A of Table A1, we conduct a formal Wilcoxon signed-rank test on whether the sample median of the 'Buy-Sell Ratio' at 9- or 1-digit ending prices is different from 1. Indeed the result suggests that the median 'Buy-Sell Ratio' across all sample AN institutions either at 9-ending or 1-ending prices is different from 1. Yet, when we further break our sample institutions based on their total dollar trading volume and compare those ranked in the top and bottom 33%, surprisingly, the degree of 'left-digit' bias seems to be more severe in 'large' institutions. We leave this puzzling result to future studies.

A2. Difference in the ‘Round-number’ Bias across Stocks Sorted by Retail Preference

As discussed in hypotheses development section (Chap. 2), departing from some prior rationales on the price clustering phenomenon (Harris, 1991; Christie and Schultz, 1994), our main conjecture is that it is mostly caused by the inherent cognitive process of human beings. An ideal testing sample shall be a large sample of trade execution and order submission data from a group of retail investors with rich information on their demographic features (e.g. Grinblatt and Keloharju, 2002). Due to the current unavailability of such type of data set, here building on a rich strand of literature on the link between specific characters of individual stocks and the preference of individual investors (e.g. Kumar, 2009), we conduct a cross-sectional analysis in this section. Specifically, following Han and Kumar (2013), we choose the idiosyncratic daily return volatility (IVOL), the idiosyncratic daily return skewness (ISKEW), the most recent institutional ownership (IO) and the closing price of the previous month (PRC) as the proxies of the retail investor’s preference. For each calendar month, we sort our sample stocks (in the TAQ sample) into deciles based on those 4 measures estimated as of the previous month and compute the EW-average of the ‘round ratio (N)’ at two prominent ‘round-number’ prices: ‘.00’ and ‘.50’. As shown in Table A2 below, in general, the average monthly ‘round ratio (N)’ of stocks that ranked in the top deciles by ‘IVOL’, ‘ISKEW’ and in the bottom decile by ‘IO’ tend to be larger than those stocks in the opposite spectrum, suggesting that the degree of price clustering might be correlated with those characters shown to be related to retail investors’ preference. Yet, the common t-test fails to reject the null (being equal between the top and bottom decile) for certain characters, as we cannot directly separate retail trades from the overall TAQ data. Another interesting finding in Table A2 is that the ‘round ratio (N)’ at ‘integer’ prices is higher for high-priced stocks, comparing with low-priced stocks. This result can be accounted by the ‘price-solution’ hypothesis in Ball et al. (1985) and Harris (1991) that for high-priced stocks, investors tend to rely on a coarser set of price intervals to reduce the negotiation costs and reach to an agreed price sooner.

A3. Robustness Check after Incorporating Competing Theories on 'Price Clustering'

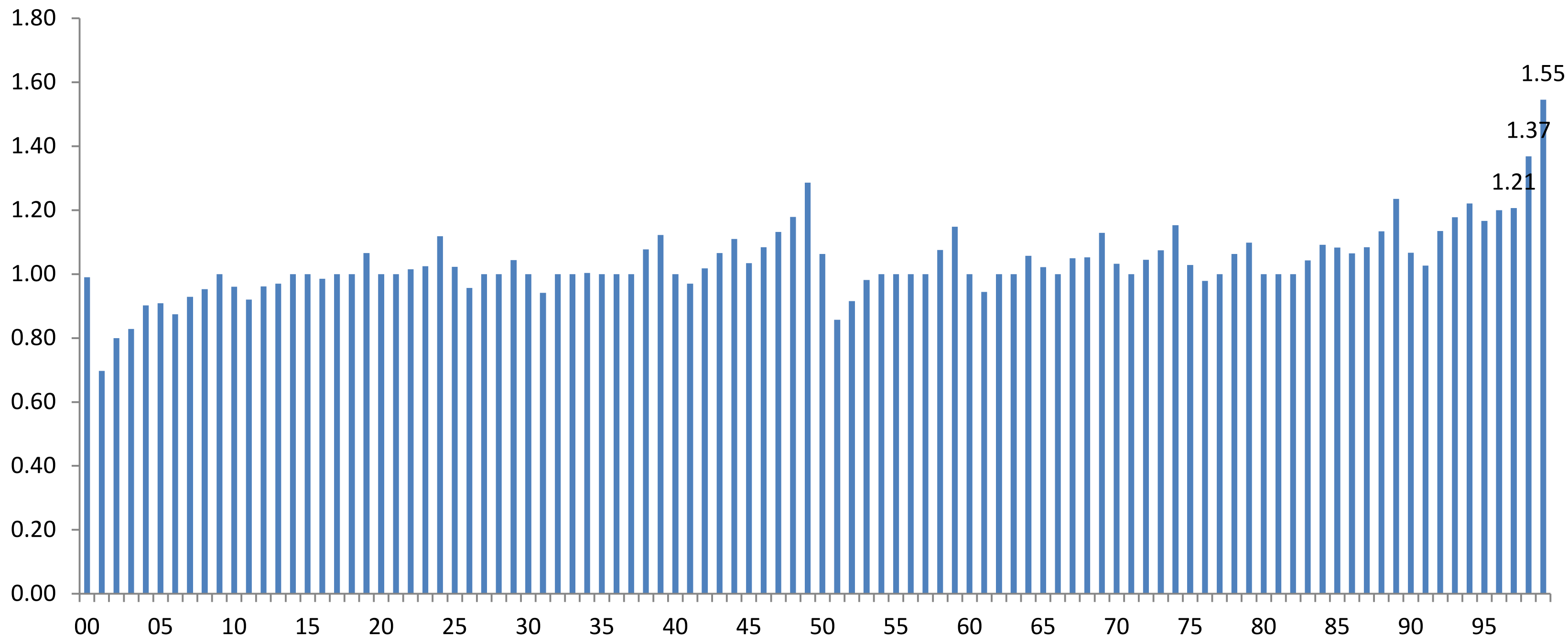
In this section, we conduct a set of regression analysis by incorporating some stock-level variables, in order to test whether our 2nd set of hypotheses regarding the role of investors' sophistication on the 'round-number' bias still hold after controlling for some well-known factors that have been shown to be related with price clustering phenomenon. In particular, as discussed in Sec.2, so far, existing literature has offered 3 main explanations for the price clustering during the transaction process. For instance, in addition to the 'price resolution' hypothesis proposed by Ball et al. (1985), where they argue that the degree of valuation uncertainty leads to the deviation from the uniform distribution, Harris (1991) adds that in a dealer-driven market, traders can use a coarse set of prices to reduce negotiation time and opportunity cost. Other scholars conjecture that the specific arrangement market arrangement may also lead to different types of price clustering, e.g. the auction market as NYSE versus the dealer market as NASDAQ (Grossman et al., 1997). Hence, building on these prior studies, here in this section, we borrow the regression specification as used by Ikenberry and Weston (2008) to further test our H2 still hold after we control for those known determinants that are related to the degree of price clustering across different stocks.

Specifically, we will conduct the following regression analysis at stock-year level (shown in Table A3) and cluster the standard error for each decile group sorted by the most recent institutional ownership level as downloaded from the Thomson Reuters 13-f data set. For the panel regression analysis in Table A3, first, we compute the 'round ratio' by number of trades executed around a' rounded penny' price for a sample stock within a calendar year and scale it by the total number of trades executed for the same stock within the same year. Then we match a set of 7 control variables and perform the pooled regression analysis. In column (3), we pool the sample stock-year observations from the AN and TAQ sample (from May, 2001 to Sep. 2011) and conduct the regression with the 'Indicator' equals 1 for the AN sample observations. The significance of the interaction terms reinforces our 2nd set of hypotheses: investors with higher level of sophistication in terms of trading skills tend to exhibit less degree of 'round-number' bias in their trading activities.

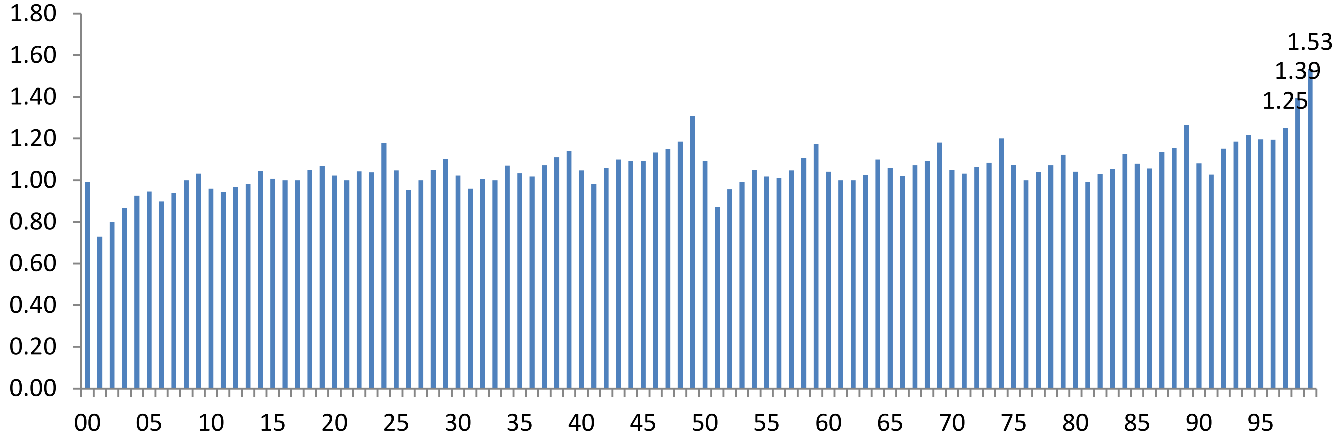
Figure A1: The left-digit bias (medians of 'Buy-Sell Ratio') of the AN sample institution (2001-2011)

Following BHJ (2012), we first count all 'valid' trades executed at the same 'rounded penny' price for a given client within a calendar year, and separately for buys and sells. Here, 'rounded penny' prices are defined as the first two post-decimal digits after directly rounding the transaction price to the nearest penny, e.g. 'X.01', 'X.02' and etc. 'Valid' trades refer to the sample Abel Noser trades that can be matched to the 'PERMNO' code of the common stocks listed in NYSE/AMEX/NASDAQ from the CRSP daily stock file. Other stock filters as stated in Sec.3. Then, we compute the 'Buy-Sell Ratio' which is the number of buy trades executed at a 'rounded penny' price (e.g. 'X.02') over the total number of all buy and sell trades at the same 'rounded penny' price for a given 'clientcode' within the same calendar year. In Pane B & C, we separate all sample institutions into two subgroups: those with in-sample total dollar trading volume is above (below) the top (bottom) 67^{th} (33^{th}) percentile are labelled as 'Big' ('Small') institutions. Finally we report the median of the 'Buy-Sell Ratio' across all 'clientcode-year' pairs at 100 different 'rounded penny' prices, from 'X.00' to 'X.99', in Panel A below. And the medians of the subsample that consists of 'Big' or 'Small' institutions are plotted in Panel B and C, respectively. Here, we only included trades submitted by institutions with client type being 1 or 2, i.e. plan sponsors and investment managers. Sample period starts from Feb. 1, 2001 for NYSE & AMEX and Apr. 2 for NASDAQ stocks and ends at Sep. 30, 2011.

Panel A: 'Left-digit bias' (All AN institutions, 2001-2011)



Panel B: 'Left-digit bias' ('Large' institutions, 2001-2011)



Panel C: 'Left-digit bias' ('Small' institutions, 2001-2011)

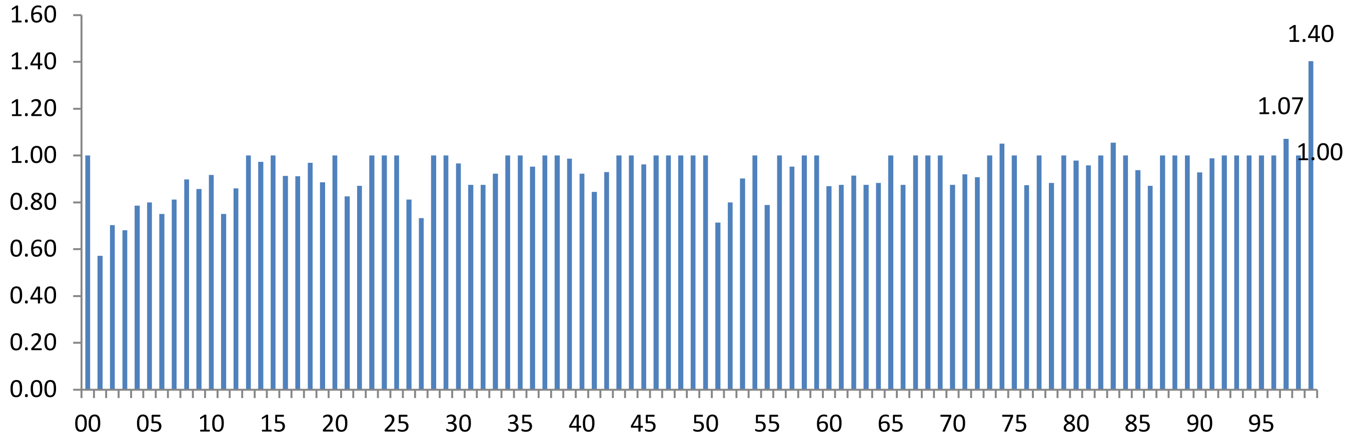


Table A1: Tests for the median of the annual 'Buy-Sell ratio' of the 9-digit (& 1-digit) ending prices (AN sample, 2001-2011)

Panel A: The Abel Noser sample (2001-2011)

H_0: $Median\ of\ \ 'Buy - Sell\ Ratio'\ (".X9") = 1$, where $X\ in\ (0, 1, \ldots, 9)$

Statistic tests	**Testing statistic**	**p-value**
Wilcoxon signed-rank	2035.5	<.0001

H_0: $Median\ of\ \ 'Buy - Sell\ Ratio'\ (".X1") = 1$, where $X\ in\ (0, 1, \ldots, 9)$

Statistic tests	**Testing statistic**	**p-value**
Wilcoxon signed-rank	-1364.0	<.0001

Panel B: Wilcoxon rank-sum test for the difference in the 'Buy-Sell ratio' between 'Big' and 'Small' institutions (AN sample, 2001-2011)

H_0: $Median\ of\ \ 'Buy - Sell\ Ratio'\ (".X9"|'Big') = Median\ of\ \ 'Buy - Sell\ Ratio'\ (".X9"|'Small')$, where $X, Y\ in\ (0, 1, \ldots, 9)$

Statistic tests	**Testing statistic**	**p-value**
Wilcoxon signed-rank	38.29	<0.001

H_0: $Median\ of\ \ 'Buy - Sell\ Ratio'\ (".X1"|'Big') = Median\ of\ \ 'Buy - Sell\ Ratio'\ (".X1"|'Small')$, where $X, Y\ in\ (0, 1, \ldots, 9)$

Statistic tests	**Statistic tests**	**Statistic tests**
Wilcoxon signed-rank	28.95	<0.001

* Here, in Panel A, we report the Wilcoxon sign-rank test for the null hypothesis that the annual 'Buy-Sell Ratio' at those 9-digit (1-digit) ending prices prices (e.g. 'X.09', 'X.19', ... , 'X.99') of the median among sample institutions in Abel Noser sample. And in Panel B, we perform the Wilcoxon test that compares the sample median of the 'Buy-Sell Ratio' of the subsample between those 'Big' and 'Small' institutions, based on the tercile sorting. Here, we only include trades that can be matched to the 'PERMNO' code from the CRSP database with the SHRCD equal to 10 or 11 and EXCHCD as 1, 2 or 3 and submitted by institutions with client type being 1 or 2, i.e. plan sponsors or investment managers.

Table A2: Tests for the difference of the 'round ratio' at two 'round-number' prices ('.00' and '.50') across stocks sorted by retail preference (with the TAQ sample, 2001-2014)

Panel A:	**Sorted by the daily return 'IVOL' in the previous month**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**D10-D1**
'.00'	*3.4%*	*3.5%*	*3.6%*	*3.7%*	*3.8%*	*3.8%*	*3.7%*	*3.7%*	*3.6%*	*3.6%*	*0.2%*
											(1.46)
'.50'	*2.4%*	*2.4%*	*2.5%*	*2.6%*	*2.6%*	*2.6%*	*2.6%*	*2.6%*	*2.6%*	*2.7%*	*0.3%*
											(3.15)
Panel B:	**Sorted by the daily return 'ISKEW' in the previous month**										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**D10-D1**
'.00'	*3.5%*	*3.6%*	*3.7%*	*3.7%*	*3.7%*	*3.7%*	*3.7%*	*3.6%*	*3.6%*	*3.6%*	*0.1%*
											(0.91)
'.50'	*2.4%*	*2.6%*	*2.6%*	*2.5%*	*2.6%*	*2.6%*	*2.6%*	*2.5%*	*2.5%*	*2.5%*	*0.1%*
											(0.98)
Panel C:	**Sorted by the lagged month-ending price (PRC)**										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**D10-D1**
'.00'	*2.2%*	*3.0%*	*3.5%*	*3.8%*	*4.1%*	*4.2%*	*3.9%*	*3.6%*	*3.3%*	*3.7%*	*1.3%*
											(9.52)
'.50'	*2.1%*	*2.3%*	*2.5%*	*2.7%*	*2.9%*	*2.9%*	*2.7%*	*2.5%*	*2.3%*	*2.2%*	*0.1%*
											(1.55)
Panel D:	**Sorted by the latest quarter-end 'IO'**										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**D10-D1**
'.00'	*6.0%*	*5.4%*	*5.0%*	*4.2%*	*3.5%*	*3.0%*	*2.8%*	*2.7%*	*2.7%*	*2.7%*	*-3.3%*
											(-16.5)
'.50'	*4.1%*	*3.7%*	*3.4%*	*2.9%*	*2.4%*	*2.1%*	*2.0%*	*1.9%*	*1.9%*	*1.9%*	*-2.2%*
											(-17.3)

Panel D:	**'Lottery'**	**'Non-lottery'**	**'Others'**	**'Lottery'- 'Non-lottery'**	
'.00'	*3.4%*	*3.5%*	*3.7%*	*-0.1%*	*(-0.89)*
'.50'	*2.5%*	*2.4%*	*2.6%*	*0.1%*	*(1.67)*

* $p < 0.05$, ** $p < 0.01$

Following Han and Kumar (2013), we compute the idiosyncratic volatility ('IVOL') using the daily returns in the previous month (*t -1*), based on the Fama-French 3-factor model (Panel A). Also, based on the same estimation period, we compute the idiosyncratic skewness ('ISKEW') measure using the model proposed by Harvey and Siddique (2000) (Panel B). And we match the closing price of the previous month to each sample stock and sort them into deciles (Panel C). And finally, we match the institutional ownership ('IO') (as the percentage of a stock's total shares outstanding) to the stock-month-price observation within 3-month window, using the Thomson Reuters' 13-F data set.

For each calendar month, we sort all sample stocks into deciles, based on one of these 4 stock characteristics. The stock-level 'round ratio' at each 'rounded penny' price is computed as follows: first for each sample transaction from the TAQ sample, we round the transaction price to the nearest penny, e.g. 'X.00', 'X.50'; then we aggregate them and count the number of trades at a given 'rounded penny' price (e.g. 'X.50') and scale by the total number of transactions for a given stock within the same calendar month. We conduct the Student's *t*-test by comparing the time-series average of the 'round ratio' at either '.00' or '.50' between the top and bottom decile and denote the *t-statistic* in the last column. And in Panel D, for each calendar month, we sort all sample stocks into 3 groups: those that are ranked below (or above) the median stock price, above (or below) the median 'IVOL' and 'ISKEW' (in total 8 groups). And those ranked in the bottom (top) group based on the PRC, and in the top (bottom) group based on 'IVOL' and 'ISKEW' independently are labelled as 'Lottery' ('Non-lottery') stocks, as in Table II of Kumar (2009). And the remaining stocks are labelled as 'Others'. Here, we only include trades that can be matched to the 'PERMNO' code from the CRSP with the SHRCD equal to 10 or 11 and EXCHCD as 1 or 3 within the normal trading hour, i.e. from 9:30 a.m. to 4:00 p.m. in Eastern Time. The sample period starts from Feb. 2001 to Dec.2014.

Table A3: Panel regressions of the difference in the degree of the 'round-number' bias across different subsamples (2001-2011)

Dependent Variable: 'Round Ratio' (%)	(1) TAQ sample	(2) AN sample	(3) AN vs TAQ
Intercept	0.81***	0.87***	0.83***
t-statistic	*39.45*	*63.97*	*52.77*
D_{00}	4.15***	3.73***	3.89***
t-statistic	*12.66*	*19.09*	*16.05*
D_{50}	2.90***	2.37***	2.71***
t-statistic	*12.11*	*16.82*	*15.35*
D_{X0}	1.69***	1.41***	1.64***
t-statistic	*24.99*	*25.93*	*30.73*
D_{X5}	1.54***	1.31***	1.49***
t-statistic	*24.88*	*33.48*	*31.64*
Intercept× Indicator			0.05***
t-statistic			*19.79*
D_{00}× Indicator			-0.16*
t-statistic			*-1.95*
D_{50}× Indicator			-0.34***
t-statistic			*-6.37*
D_{X0}× Indicator			-0.23***
t-statistic			*-33.18*
D_{05}× Indicator			-0.18***
t-statistic			*-19.19*
Std. Errors Clustering	IO-level	IO-level	IO-level
Control Variables	Yes	Yes	Yes
Adjusted R^2	67%	13%	22%
N	4,978,200	4,559,600	8,467,200

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

p-value in Italic.

In column (1) to (3), we first round the transaction price to the penny level. Then we count the number of trades executed at a given rounded price (e.g. '.01') for a sample stock within a calendar year and divide it by the total number of transactions for the same stock within the same year (i.e. the 'round ratio'). Next, we regress the stock-level 'round ratios' on a series of 'roundness' indicators (e.g. D_{X0} equals 1 for the 'rounded penny' price '.X0' where X is an integer ranging from 1 to 9, excluding 5) and the interaction terms (which equal 1 if the 'round ratio' is computed based on trades from the AN sample), as in Equation (2) in KLZ (2015). We conduct the pooled regression with a series of 7 stock-level characteristic variables and cluster the standard error at the institutional ownership level (sorted into deciles for each sample year). In column (3) the interaction term ('*Indicator*') equals 1 if the transaction is from the Abel Noser sample.

Following Ikenberry and Weston (2007), we include a set of 7 stock-level control variables in our panel regression:

Avg_BA (average bid-ask spread from Jul. to Dec. in the past year), logMV (natural logarithm of the market capitalization as of Dec. in the past year), 'ISKEW' (based on daily returns from Jul. to Dec. in the past year using the Harvey-Siddique model), Avg_Trade_Size (the ratio between the total shares traded and the total number of trades for a sample stock in the current year), Volatility is the standard deviation of the daily returns from Jul. to Dec. in the past year and a dummy variable for NASDAQ listed stocks. Each year, we sort our sample stocks into deciles based on the institutional ownership level (computed using the Thomson Reuters 13-F data set as of Dec. in the past year) and computed the t-statistics based on the standard errors clustered at each decile.