The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXIX, NO. 4 • AUGUST 2024

A (Sub)penny for Your Thoughts: Tracking Retail Investor Activity in TAQ

BRAD M. BARBER, XING HUANG, PHILIPPE JORION, TERRANCE ODEAN, and CHRISTOPHER SCHWARZ*

ABSTRACT

We placed 85,000 retail trades in six retail brokerage accounts from December 2021 to June 2022 to validate the Boehmer et al. algorithm, which uses subpenny trade prices to identify and sign retail trades. The algorithm identifies 35% of our trades as retail, incorrectly signs 28% of identified trades, and yields uninformative order imbalance measures for 30% of stocks. We modify the algorithm by signing trades using the quoted spread midpoints. The quote midpoint method does not affect identification rates but reduces the signing error rates to 5% and provides informative order imbalance measures for all stocks.

BOEHMER ET AL. (2021, BJZZ), PROPOSE AN ALGORITHM that identifies as retail trades those that are executed off exchanges, typically by wholesalers such as Citadel or Virtu, and receive subpenny price improvement. The key idea relies on Regulation NMS (National Market System), adopted in 2005 by the Securities and Exchange Commission (SEC). Rule 612 prohibits the display of quotes with decimals below one cent (the "subpenny" rule) while still allowing execution at prices with fractional cents. Wholesalers typically offer

*Brad Barber is at the Graduate School of Management, UC Davis. Xing Huang is at the Olin Business School, Washington University in St. Louis. Philippe Jorion is at The Paul Merage School of Business at the University of California at Irvine. Terrance Odean is at the Haas School of Business, University of California, Berkeley. Christopher Schwarz is at The Paul Merage School of Business at the University of California at Irvine. We appreciate the comments of Sida Li, Yuqing Yang, and seminar participants at Baruch College and Humboldt University. We are grateful for the comments of two reviewers, the associate editor, and Jonathan Lewellen (the editor). We wish to thank Alyssa Moncrief for research assistance. None of the authors has conflicts of interest to disclose. Barber serves on the advisory board of Vert Asset Management. Odean is a member of the academic advisory board of Matson Money, an advisory editor to the Financial Planning Review, a member of the academic advisory board of *The Journal of Investment Consulting* and was previously a member of the academic advisory board of Russell Investments.

Correspondence: Christopher Schwarz, The Paul Merage School of Business at the University of California at Irvine; 4293 Pereira Drive, Irvine, CA, 92697-3125. E-mail: cschwarz@uci.edu.

DOI: 10.1111/jofi.13334

price improvement to a large fraction of retail trades that are routed to them. Such price improvement is often a subpenny. Institutional orders, in contrast, are typically not executed by these wholesalers. Accordingly, the BJZZ algorithm assumes that institutions represent only a small part of the trades that execute at a subpenny off exchanges. In addition, the algorithm does not identify off-exchange trades that execute at subpennies between \$0.004 and \$0.006 as retail, because "…trades at or near the half-penny are likely to be from institutions" (BJZZ, p. 2251).

The BJZZ algorithm also uses the subpenny digit to sign a trade as a buy or sell. Specifically, if the subpenny amount is less (more) than \$0.004 (\$0.006), the trade is labeled as a sell (buy). BJZZ validate the algorithm using a sample of 117 NASDAQ stocks in 2010 and study the algorithm from 2010 to 2015. The algorithm is the best tool available to identify retail trades in the NYSE Trade and Quote (TAQ) data set and therefore is increasingly used in academic studies to measure the intensity of retail trading and retail investor order imbalance.[1] Most of these studies, however, are set in more recent periods with substantially different market structures (e.g., no commissions).[2]

To provide guidance to scholars relying on this method to study retail trading, we ran a trading experiment from December 2021 to June 2022 to evaluate the BJZZ algorithm's ability to identify and accurately sign trades. We call the rate at which the algorithm correctly identifies our retail trades as the *identification rate* and the rate at which the algorithm correctly signs our trades as a buy or sell, conditional on the trade being identified, as the *accuracy rate*.[3] We also compare BJZZ's signing method to the Lee and Ready (1991) quote midpoint signing method. To implement the quote midpoint method, we sign a trade as a buy (sell) if the execution price is greater (less) than the quote midpoint, but we do not sign trades that execute between 40% and 60% of the National Best Bid or Offer, or NBBO. We find that the two methods have similar identification rates. However, we further find that the use of subpenny digits to sign trades leads to low accuracy rates for stocks with spreads greater than a penny. In contrast, we find that the quote midpoint method leads to high, homogeneous accuracy rates across all stocks. We therefore recommend that researchers use the quote midpoint method to sign retail trades identified through subpenny pricing.

Figure 1 illustrates how the BJZZ subpenny digit and the quote midpoint methods sign trades for stocks that face a spread of one, two, or three cents.

[1] Papers that analyze retail volume include Blankespoor et al. (2019), Bushee, Cedergren, and Michels (2020), Bonsall, Green, and Muller (2020), Farrell et al. (2022), Israeli, Kasznik, and Sridharan (2022), and Guest (2021). Papers that analyze retail order imbalance include Barber, Lin, and Odean (Forthcoming), Farrell et al. (2022), and Bradley, Jame, and Williams (2022).

[2] After the BJZZ sample period, the SEC Tick Size Pilot (TSP), which took place for about two years beginning in October 2016, limited quotes and trades in two test groups (G2/G3) to \$0.05 increments with certain exceptions. There is a dramatic drop in subpenny price trades reported on TAQ for stocks in these two groups, so we recommend that scholars generally exclude stocks in these groups during the TSP period. See https://www.finra.org/rules-guidance/key-topics/tick-size-pilot-program/data-collection-securities-and-pilot-securities-files for details.

[3] Throughout the paper, we refer to the conditional accuracy rate as the "accuracy rate" for the sake of brevity.

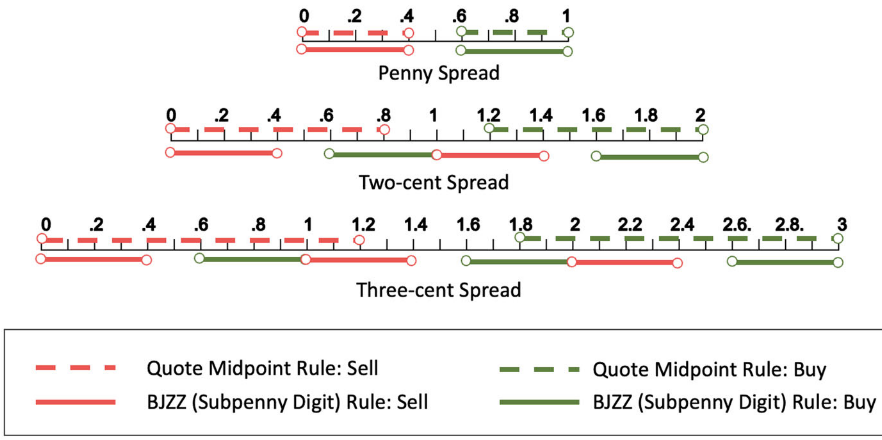

Figure 1. Trade signing illustration: BJZZ (subpenny digit) rule vs. quote midpoint rule. In this figure, we illustrate the signing of both the BJZZ (subpenny digit) rule and the quote midpoint rule by using three examples: penny spread, two-cent spread, and three-cent spread. (Color figure can be viewed at wileyonlinelibrary.com)

When spreads are one penny, the two algorithms are identical in how they sign the direction of trade. When spreads are two cents, the subpenny digit method incorrectly signs as buys those trades that execute just *below* the quote midpoint but have a subpenny digit greater than \$0.006; similarly, the method incorrectly signs as sells those trades that execute just *above* the midpoint but have a subpenny digit less than \$0.004. The subpenny digit method therefore creates classification errors if trades around the midpoint are common, unlike the quote midpoint rule. For spreads of three cents, misclassifications now occur over a wider range, so the potential problem becomes more pronounced as the spread increases. This issue is the main driver of low accuracy rates for the BJZZ algorithm.[4]

In our trading experiment, we executed more than 85,000 trades over a six-month period in six retail brokerage accounts at five different brokers beginning in December 2021. Our main sample focuses on 64,000 trades made in 128 stocks each day. The 128 stocks represent a stratified sample of all stocks, which allows us to draw inferences about the performance of the two algorithms for the average stock and average trade. We conduct several additional targeted trading experiments to test the robustness of our main findings.

The BJZZ algorithm identifies 35% of the trades that we placed as retail, which is likely close to results reported in BJZZ.[5] Identification rates vary systematically across stocks, with the variation economically large, ranging from

[4] The quote midpoint method signs trades that execute at subpennies between \$0.004 and \$0.006 (including \$0.005) if the trade does not execute between 40% and 60% of the NBBO.

[5] BJZZ do not report a directly comparable number but do show that 46% of trades in 117 NASDAQ stocks on retail venues in October 2010 execute at subpennies other than the half-penny. Among our trades, 42% execute at subpennies other than the half-penny.

20% to 50%. Subpenny price improvement is much more common for trades with a one-penny spread than for trades with spreads greater than a penny. As a result, identification rates drop as spreads increase. Figure 2, Panel A, shows that BJZZ identification rates are close to 40% for stocks with a spread of 5 cents or less, but drop to 28% for stocks with spreads greater than 10 cents. Identification rates using the quote midpoint method are similar since it also relies on subpenny execution to identify retail trades.

Turning to accuracy, the BJZZ algorithm signs 72% of identified trades correctly (a 28% error rate). As with identification, wider spreads are the main reason the subpenny digit fails to accurately sign trade direction. In Figure 2, Panel B, we show that the accuracy rate for the BJZZ algorithm is 93% for identified trades that face a penny spread but drops to 52% when spreads are 10 cents or greater. Note that in our sample, more stocks face a spread of 10 cents or greater than face a penny spread (34% versus 20%, as shown on the horizontal axis of Figure 2). For 30% of our stocks, we fail to reject the null hypothesis that the signing accuracy rate for the stock is 50%. In contrast, the quote midpoint method accurately signs 95% of identified trades (a 5% error rate) and the accuracy rate does not vary significantly across spread bins.

Subpenny signing is sensitive to spread size because the implicit assumption that retail trades receive minimal price improvement is often violated in our trading experiment. In Figure 3, Panel A, we present a histogram of execution prices for our sell trades relative to the spread conditional on subpenny execution.[6] Though the bulk of sell transactions execute just above the bid, a good fraction of trades executes near or even above the midpoint. In Panel B, we observe analogous patterns for buys, which tend to cluster just below the ask, but with many executions near or even below the midpoint.

To visualize the effect of these execution patterns on signing accuracy, in Panel C, we plot signing accuracy for all subpenny trades using either the subpenny digit (as in BJZZ) or the quote midpoint method. Several points are worth noting. First, the quote midpoint method clearly dominates the subpenny digit method at every trade location. Near the quote midpoint, the subpenny digit method drops below 50% accuracy because spreads quoted in even cent amounts (e.g., $0.02) will incorrectly sign executions that occur near the quote midpoint, as illustrated in Figure 1. However, accuracy drops dramatically near the quote midpoint for the quote midpoint method as well. For this reason, we recommend eliminating trades that occur between 40% and 60% of the posted NBBO when applying the quote midpoint method.

We should note that our subpenny digit accuracy level of 72% is much lower than the 98.2% reported in BJZZ. This difference can be traced to changes in the size of spreads over time. Penny spreads were much more common in 2010. Figure 4 plots the cumulative density of spreads at six-year intervals beginning in 2010. The figure shows that the frequency of penny spreads has dropped dramatically from about 80% of all off-exchange trades in 2010 to about 40% in 2022.

[6] We also exclude trades that execute at the quote midpoint if it is a half penny since neither algorithm identifies trades that execute at the quote midpoint as retail trades.

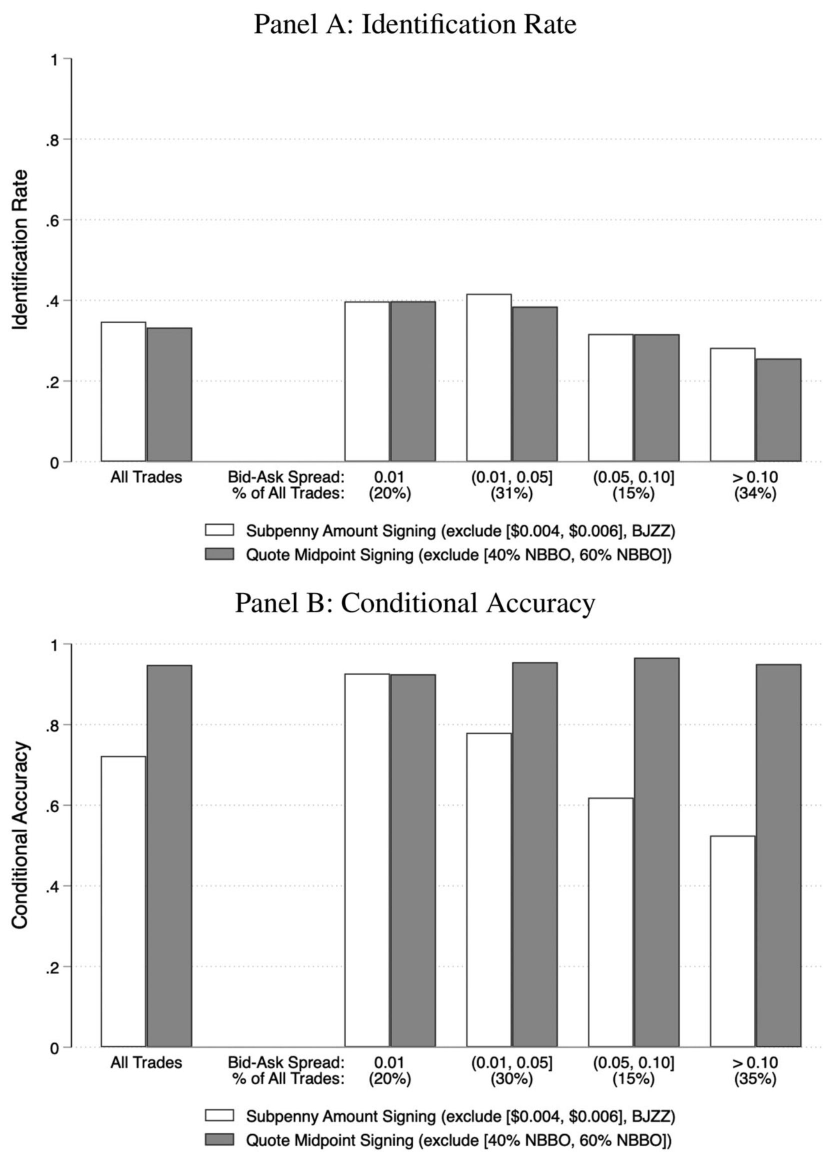

Figure 2. Identification and signing accuracy by bid-ask spread for modified signing. In this figure, we graph the identification rate (Panel A), which is the percentage of our retail trades identified as retail, and the signing accuracy (Panel B), which is the probability that the signing is accurate conditional on the trade being identified. We compare across two algorithms: (1) BJZZ: subpenny amount signing (exclude [$0.004, $0.006]); (2) quote midpoint signing (exclude [40% NBBO, 60% NBBO]). We report results for our entire sample as well as four different ranges of bid-ask spread. The identification and conditional accuracy rates at one penny are not precisely equal for the two algorithms because the small number of trades that execute outside the NBBO are identified and signed differently by the two algorithms. Under each bin, we report the percentage of our trades that fall within that range. The figures are based on our 64,164 trades of the stratified sample made between December 2021 and June 2022.

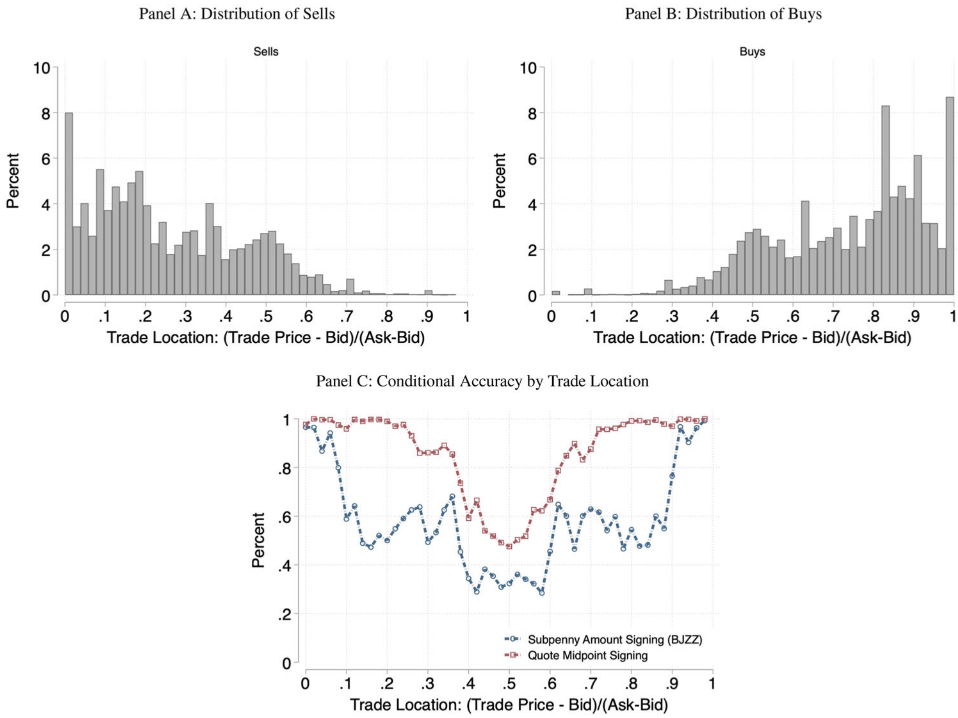

Figure 3. Distribution of buy and sell trades and conditional accuracy across trade location within spread. The figure presents the histogram of sells (Panel A) and buys (Panel B) in each region of trade location between bid and ask prices. Trade location is defined as the distance between the trade price and the bid price divided by the distance between the ask price and the bid price. Panel C presents the accuracy rate conditional on being identified for each region of trade location for the subpenny amount (BJZZ) algorithm and the quote midpoint algorithm. The figures are based on our 64,164 trades of the stratified sample made between December 2021 and June 2022. (Color figure can be viewed at wileyonlinelibrary.com)

Since accuracy rates decline as spreads increase, the accuracy of the BJZZ subpenny digit method to sign trades has become less reliable in recent years. In sharp contrast, the quote midpoint method's accuracy rate of 95% is similar across spread levels (see Figure 2, Panel B). This is an important improvement for scholars who seek to measure retail order imbalance using off-exchange subpenny trades. As one example, we show that the BJZZ algorithm yields little change in retail order imbalance during the infamous GameStop episode in January 2021, while the quote midpoint method shows a dramatic increase in retail buying (see Internet Appendix Section I and Figure IA.1 [7]).

While our suggested signing method improves accuracy, it does not improve the identification rate. Our trading experiment suggests that false negative identifications are common (65% of our retail trades are not classified as retail). However, an algorithm designed to identify retail trades can yield false

[7] The Internet Appendix may be found in the online version of this article.

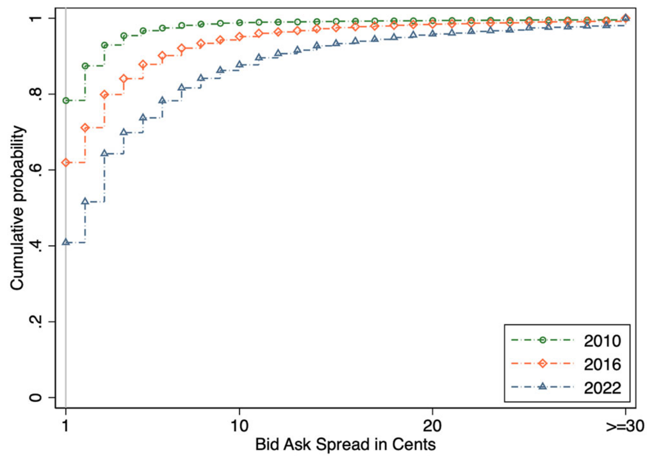

Figure 4. Cumulative distribution function (CDF) plot for bid ask spreads of TAQ universe over time. In this figure, we graph the CDF of the bid-ask spread across all trades from Exchange D recorded in TAQ. For each year, we sample the 10th of each month to create the distribution, and equally weight the cumulative probability across stocks. (Color figure can be viewed at wileyonlinelibrary.com)

positives as well (i.e., classify nonretail or institutional trades as retail). BJZZ suggest that one reason to exclude trades near the quote midpoint is because these trades are more likely to be institutional. Unfortunately, we cannot say anything definitive about the rate of false positives, which would require a representative sample of institutional trades to identify the frequency of institutional subpenny price execution off exchanges. Battalio et al. (2022) argue that a large proportion of nonretail trades executed by wholesalers execute at subpennies other than a half penny, so the possibility that false positives obscure the signal regarding the intensity or direction of retail trade is an important caveat to scholars using the algorithms we study to identify and sign retail trades.

Our estimates of identification and accuracy rates are based on a sample of approximately $100 trades in whole shares that we placed in six brokerage accounts at five brokers in 2021 and 2022. These estimates might vary across different trade sizes, brokers, and time. In a limited sample of trades of about $1,000, we find similar results. At larger trade sizes, however, price improvement relative to the spread may be less generous, which might generate higher identification and accuracy rates for the subpenny digit method (although the quote midpoint method should still perform well when trades receive minimal price improvement).

Price improvement varies systematically across retail brokers (Levy (2022), Schwarz et al. (forthcoming)). Thus, our estimates of identification and accuracy rates assume that the brokers that we use (E*Trade, Fidelity, Interactive Brokers, Robinhood, and TD Ameritrade) are representative of the trades placed across all retail brokers. This is a reasonable assumption since the brokers we use execute a large fraction of total retail trading.[8] Finally, the prevalence of penny spreads was much greater during the period analyzed by BJZZ, so the quote midpoint method would offer somewhat more modest improvements in signing accuracy in these earlier periods.[9] Even with these caveats, we conclude that the midpoint quote rule will dominate the subpenny digit rule in assessing the direction of orders because of the inherent limitations of the digit to sign trades that face a spread of more than a penny and receive more than de minimis price improvement.

Overall, our results have important implications for those who rely on off-exchange subpenny prices to estimate the frequency of retail trading and retail order imbalance. Hypothesis tests about the frequency of retail trading face a joint testing problem. The null hypothesis that two stocks have equal rates of retail trading is a joint test of two hypotheses: first, that retail trading is equal in the two stocks, and second, that the ability of the algorithm to identify retail trading is the same in the two stocks. Because subpenny execution is more common for trades that face lower spreads, variation in spreads across stocks or over time will cause the frequency of identified retail trades to vary. Ignoring this variation in identification rates can lead to spurious inferences regarding the intensity of retail trading.

Hypothesis tests about retail order imbalance face a similar joint testing problem. The null hypothesis that two stocks have equal retail order imbalance is a joint test of two hypotheses: first, that retail order imbalance is equal in the two stocks, and second, that the ability of the algorithm to estimate retail order imbalance is the same for the two stocks. Fortunately, we show that the quote midpoint method that we propose substantially reduces the joint testing problem by breaking the link between spreads and signing accuracy rates.

The remainder of the paper is organized as follows. In Section I, we describe the experiment. In Section II, we summarize results. In Section III, we discuss the issue of false positives and implicit assumptions about excluded trades. Finally, Section IV concludes.

I. Description of the Experiment

In this section, we provide a brief description of the trading experiment. Details are provided in Internet Appendix Section II. We made about 85,000 trades between December 2021 and June 2022 in six brokerage accounts at five

[8] Retail trading activity can be compared across brokers that report "daily average trading" (DAT) statistics. DAT represents average trades per day that generate commissions, fees, or payment for order flow, and now also includes commission-free trades. In the first half of 2022, reported DATs were 4.7MM, 3.0MM, 2.3MM, 2.1MM, and 0.9MM for TD Ameritrade, Fidelity, Robinhood, IBKR, and E*Trade, respectively, for a total of 13.0MM. Our sample therefore accounts for a large fraction (87%) of the seven DAT-reporting brokers (Schwab is at 1.7MM and TradeStation at 0.2MM.)

[9] In Internet Appendix Figure IA.2, we show that the correlation of daily retail order imbalance for the two methods across stocks averages 86% from 2010 to 2015. Post-2015, the mean daily correlation drops to 67%. Correlations are lower for stocks with more than a penny spread in both periods. These correlations do not capture the attenuation bias caused by the lower accuracy rates, which we describe in Internet Appendix Section II and Figure IA.7.

Table I

Descriptive Statistics for Our Trades

This table reports summary statistics for our 64,164 trades of the stratified sample made between December 2021 and June 2022. Price is the executed price. Bid-Ask Spread is the quoted bid-ask spread just prior to trading. Dollar Size is the dollar amount traded whereas Trade Size is the number of shares traded. Split Trades is one if our trade was split into multiple trades. Return is the daily return of the stock. Abnormal Volume is the day's volume divided by the prior 21-day average. Abnormal Volatility is the absolute daily return divided by the prior 21-day average.

	Mean	Std. Dev.	p10	p25	Median	p75	p90
Price	\$54.69	\$94.58	\$2.43	\$6.24	\$17.94	\$60.48	\$164.74
Bid-Ask Spread	\$0.17	\$0.41	\$0.01	\$0.02	\$0.05	\$0.15	\$0.35
Dollar Size	\$118.14	\$73.02	\$86.28	\$97.06	\$100.02	\$106.04	\$165.12
Trade Size (Shares)	14.05	20.56	1.00	2.00	5.00	16.00	40.00
Split Trades (%)	1.80%	13.31%					
Return	0.00	0.04	−0.04	−0.02	0.00	0.01	0.04
Abnormal Volume	1.10	2.50	0.36	0.59	0.83	1.16	1.65
Abnormal Volatility	1.07	1.10	0.12	0.37	0.81	1.46	2.24

retail brokers: Robinhood (RH), E*Trade (ET), Fidelity (FD), TD Ameritrade (TD), and two accounts at Interactive Brokers (IBKR, "Pro" and commission-free "Lite").

We traded 128 stocks each day by selecting one stock from each of 128 bins, which were created based on market capitalization, daily volatility, daily share turnover, and stock price. We targeted a trade size of \$100 and traded full shares only (rounding the shares traded to make the trade size as close to \$100 as possible). In robustness checks, we conduct auxiliary experiments that trade in round lots, larger amounts (\$1,000 or up to \$5,000), or targeted stocks (stocks popular with retail investors, megacap stocks with the majority of overall trading volume, and top movers based on overnight returns).

We use the trade execution time and price reported by the broker to match our trades to the TAQ database, which provides the exchange code, the execution time on the exchange, and the prevailing NBBO at the time of the trade.[10] Our main sample consists of more than 64,000 trades with a total dollar value of \$7.6 million. The auxiliary experiments add about 20,000 trades worth about \$8 million.

Table I presents summary statistics for our main sample. The mean (median) stock price is \$55 (\$18). Importantly, the average NBBO spread is 17 cents and the median is 5 cents. More than 80% of trades in our sample

[10] We use the WRDS Consolidated Trades (WCT) files to identify off-exchange trades (exchange code "D") that execute at a subpenny amount. Since 2017, WRDS calculates the NBBO in the nanosecond prior to the trade. Only 125 of our trades execute outside the NBBO and we include them in our analysis. For the trades outside the NBBO that are identified as retail, the BJZZ subpenny digit has an accuracy rate of 77% and the quote midpoint method has an accuracy rate of 65%. Since the NBBO is likely estimated with error in these cases, it would be reasonable to use the BJZZ algorithm for trades outside the WRDS-calculated NBBO and when the NBBO spread is a penny.

face a spread greater than one penny.[11] The trade size averages $118 because we target $100 trades while trading round shares. Few trades are split (1.8%). Because we begin trading at TD and RH first, 37% of our trades are from these two brokers, but more than 1,000 trades are placed through each of the six brokerage accounts.

II. Results

A. Identification Rates

Table II summarizes the ability of the BJZZ algorithm to identify retail trades. The fraction of total trades with price improvement (PI) exceeds 90%; 61% of trades receive subpenny price improvement, and 41% of trades execute at a subpenny other than $0.005 (i.e., 20% of trades execute at $0.005). BJZZ reports similar estimates for a NASDAQ TRF sample in 2010. BJZZ implement an exclusion "donut" of [$0.004, $0.006] because "…trades at or near the half-penny are likely to be from institutions" (BJZZ, p. 2251). When we add this filter to our trades, the BJZZ algorithm identifies 35% of our sample trades as retail. In the last column, we find that the quote midpoint method, which excludes trades within 40% and 60% of the NBBO, has a similar identification rate.

Each stock is a random draw from one of the 128 bins generated to construct the stratified sample. Since we have a roughly similar number of trades in each stock, equal-weighted statistics provide results for the average stock.[12] To obtain statistics for the average trade, we weight each bin by either the total dollar volume of trade in the bin (VW) or the total off-exchange volume for the bin (ExDW). In the second and third rows, we show that results for the average trade are quite similar to those for the average stock (first row). The fourth and fifth rows show that identification rates are similar for buys and sells. In the last six rows, we observe differences in identification rates across brokers, ranging from 11% for trades placed at IB Pro to 42% for trades placed at E*Trade for the BJZZ algorithm, and 14% at IB Pro to 47% at E*Trade for the quote midpoint method. These differences arise because price execution varies across brokers (Schwarz et al. (forthcoming)).

[11] The fact that about 20% of trades in our sample face a one-penny spread but during our sample period about 40% of trades have a one-penny spread can be traced to the timing of our trades, which is uniform across the day. If penny spreads are more common during periods of high volume (i.e., spreads are lower when there is more liquidity in the market), then uniform spacing of trades will generate a lower percentage of one-penny spreads. Of course, ex ante we are not able to sample in proportion to contemporaneous trading volume. Internet Appendix Figure IA.3 shows that the proportion of penny spreads is similar for both all stocks and our stocks if we only analyze trades at the same time of the day as our trades.

[12] The 128 bins in our main sample do not have exactly the same number of trades. When we create sample weights based on the number of trades in each bin, all statistics in the first row of Table II are within 0.24% of those presented. For simplicity, we weight each trade equally.

Table II

Identification Rate Comparison: Subpenny Amount (BJZZ) vs. Quote Midpoint Signing

This table displays the proportion of our 64,164 trades in the stratified sample made between December 2021 and June 2022 that receive any price improvement (PI), subpenny PI, subpenny PI (exclude $0.005), subpenny PI (exclude [$0.004, $0.006], identified as retail trades by BJZZ algorithm), and subpenny PI (exclude [40% NBBO, 60% NBBO], identified as retail trades by quote midpoint signing algorithm). In the first row, we report results equally weighted across all trades (EW). Stocks are sampled from 128 bins to create a stratified sample. In the second row, trades are weighted by the bin's total dollar volume (VW). In the third row, trades are weighted by the bin's off-exchange dollar volume (ExDW). Volume data are from TAQ.

	Number of Trades	PI	Subpenny PI	Subpenny PI (Exclude $0.005)	Subpenny PI (Exclude [$0.004, $0.006])	Subpenny PI (Exclude [40% NBBO, 60% NBBO])
Total (EW)	64,164	91.32%	60.67%	41.38%	34.71%	33.26%
Total (VW)	64,164	93.67%	55.77%	36.93%	30.19%	28.46%
Total (ExDW)	64,164	93.63%	55.84%	36.95%	30.19%	28.49%
Sell (EW)	32,028	90.42%	61.38%	41.90%	35.14%	33.54%
Buy (EW)	32,136	92.22%	59.97%	40.86%	34.29%	32.98%
E*Trade (EW)	11,081	95.99%	68.08%	49.83%	42.18%	47.07%
Fidelity (EW)	1,045	92.63%	63.16%	49.00%	40.67%	40.96%
Interactive Brokers Pro (EW)	4,169	76.90%	23.87%	12.52%	11.20%	13.55%
Interactive Brokers Lite (EW)	1,010	63.07%	36.73%	25.94%	24.75%	25.84%
Robinhood (EW)	23,459	84.53%	57.04%	42.39%	37.03%	42.24%
TD Ameritrade (EW)	23,400	99.65%	68.29%	41.84%	33.21%	21.20%

B. Accuracy Rates

Table III presents the conditional accuracy rates for the BJZZ and quote midpoint algorithms. The BJZZ algorithm accurately signs 72% of trades as a buy or sell. The quote midpoint method accurately signs 95% of trades. The equally weighted statistics represent statistics for the average stock. In the second and third rows, we observe similar accuracy rates for the average trade (weighting stocks alternatively by the bin's total volume or the bin's off-exchange volume). In the fourth and fifth rows, we find that buys and sells have very similar accuracy rates.

In the last six rows, we find that conditional accuracy rates also vary significantly across brokers for the BJZZ algorithm, from 64% at TD Ameritrade to 88% at IB Lite. The quote midpoint rule reduces this variation substantially. Accuracy for TD Ameritrade, however, is notably lower than for the other brokers because our trades at this broker receive price improvement close to 50% (Schwarz et al. (forthcoming)).

C. Multivariate Analysis of Identification Rates

Spreads affect identification rates because actual subpenny execution drops as spreads increase. To explore whether other factors affect identification rates, we estimate a linear probability regression that predicts identification. The dependent variable is equal to one if the trade is correctly identified as retail using the BJZZ algorithm and zero otherwise. We present results in Table IV. Results for the BJZZ subpenny digit method are presented in columns (1) to (4). Analogous results for the quote midpoint method are presented in columns (5) to (8).

In column (1), we estimate a regression with indicator variables for different spread buckets (using the one-penny spread as the omitted category). For the BJZZ algorithm, we find that identification rate probabilities drop by 9.4 percentage points (ppt) when spreads are between 5 cents and 10 cents and by 12.5 ppt when spreads exceed 10 cents. The odd spread indicator takes a value of one when the spread is in odd cents. It is statistically significant but has a relatively modest negative effect on identification rates. The negative effect reflects the fact that a modestly disproportionate share of trades occurs near the midpoint quote (Figure 3) and, for stocks with odd spreads, those trades are dropped by the BJZZ algorithm because they have a subpenny increment near \$0.005. The effect is smaller in column (5) because the midpoint quote method treats stocks with odd and even spreads symmetrically.

In column (2), we consider several stock characteristics: the log of stock price, return quintiles, an extreme return indicator (top 0.5% of same-day return), a negative return indicator, quintiles of abnormal volume and volatility, and market condition variables (up market indicator and standardized VIX level). The price levels effect is important because spreads are correlated with price levels. There is weak evidence of identification being lower in up markets, but the effect is economically small. The effect of the VIX variable is statistically

Table III

Accuracy Rate Comparison: Subpenny Amount (BJZZ) vs. Quote Midpoint Signing

This table compares the conditional accuracy (i.e., the probability that the signing is accurate conditional on the trade being identified) between algorithms using subpenny amount signing (BJZZ) and quote midpoint signing. We report results for all trades taken together as well as buys and sells and each broker separately. In the first row, we report results equally weighted across all trades (EW). Stocks are sampled from 128 bins to create a stratified sample. In the second row, trades are weighted by bin's total dollar volume (VW). In the third row, trades are weighted by the bin's off-exchange dollar volume (ExDW). Volume data are from TAQ. The table is based on our 64,164 trades of the stratified sample made between December 2021 and June 2022.

Identification Method: / Signing:	Subpenny PI (Exclude [$0.004, $0.006]) Subpenny Amount (BJZZ)		Subpenny PI (Exclude [40% NBBO, 60% NBBO]) Quote Midpoint	
	Number of Trades	P(Correct | Identified)	Number of Trades	P(Correct | Identified)
Total (EW)	22,274	72.22%	21,323	94.79%
Total (VW)	19,373	63.98%	18,246	95.05%
Total (ExDW)	19,370	64.10%	18,268	95.08%
Sell (EW)	11,256	72.43%	10,733	94.65%
Buy (EW)	11,018	72.00%	10,590	94.94%
E*Trade (EW)	4,674	72.83%	5,214	97.09%
Fidelity (EW)	425	70.35%	428	98.13%
Interactive Brokers Pro (EW)	467	84.58%	562	99.65%
Interactive Brokers Lite (EW)	250	87.60%	261	99.23%
Robinhood (EW)	8,687	78.59%	9,904	99.81%
TD Ameritrade (EW)	7,771	63.60%	4,954	81.29%

Table IV

Drivers of the Identification Rate: Subpenny Amount (BJZZ) vs. Quote Midpoint Signing

The table presents a linear probability model that examines the drivers that impact the identification rate of algorithms using subpenny amount signing (BJZZ) and quote midpoint signing. Our trades are made between December 2021 and June 2022. The dependent variable is an indicator that equals one if a trade is identified as a retail trade by one of the algorithms. The independent variables include three sets of variables: (1) trade characteristics: indicators for different bid-ask spread regions ((\$0.01, \$0.05], (\$0.05, \$0.10], and greater than \$0.10) and an indicator for spreads that are odd number cents; (2) stock characteristics on the trading day: the logarithm of stock price, the return quintile group among all CRSP stocks, the extreme return dummy (equals one if the absolute return is in the top 0.5 percentile among all CRSP stocks), an indicator for negative returns, the quintile group of abnormal volume ($Vol(t)/AvgVol(t-20, t-1)$), and abnormal volatility ($AbsRet(t)/AvgAbsRet(t-20, t-1)$) among all CRSP stocks; and (3) market conditions on that day: an indicator for positive market returns and standardized VIX. Standard errors are clustered at the firm and day level and reported in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Dep Var:	Identification Indicator							
Identification Method:	Subpenny PI (exclude [\$0.004, \$0.006])				Subpenny PI (exclude [40% NBBO, 60% NBBO])			
Signing:	Subpenny Amount (BJZZ)				Quote Midpoint			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spread > \$0.01 and Spread ≤ \$0.05	0.00892		0.0195*	−0.0142	−0.0131		−0.00203	−0.0273**
	(0.010)		(0.010)	(0.013)	(0.009)		(0.009)	(0.012)
Spread > \$0.05 and Spread ≤ \$0.10	−0.0944***		−0.0731***	−0.0985***	−0.0813***		−0.0593***	−0.0686***
	(0.011)		(0.011)	(0.017)	(0.010)		(0.010)	(0.015)
Spread > \$0.10	−0.125***		−0.0922***	−0.111***	−0.140***		−0.106***	−0.115***
	(0.010)		(0.010)	(0.017)	(0.009)		(0.010)	(0.016)
Odd Spread	−0.0174***		−0.0164***	−0.0175***	−0.000439		0.000560	0.00116
	(0.005)		(0.005)	(0.005)	(0.004)		(0.004)	(0.004)
Ln(Price)		−0.0306***	−0.0138***	−0.0185		−0.0326***	−0.0145***	−0.0124
		(0.002)	(0.002)	(0.018)		(0.002)	(0.002)	(0.016)
Return Quintile		−0.00104	−0.000212	−0.000781		−0.000245	0.000691	−0.000148
		(0.002)	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)

(*Continued*)

Table IV—*Continued*

Dep Var:	Identification Indicator							
Identification Method:		Subpenny PI (exclude [$0.004, $0.006])				Subpenny PI (exclude [40% NBBO, 60% NBBO])		
Signing:		Subpenny Amount (BJZZ)				Quote Midpoint		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Extreme Return Dummy		−0.0557	−0.0460	−0.0446		−0.0603	−0.0532	−0.0308
(=1: absolute return in top 0.5 percentile)		(0.043)	(0.042)	(0.045)		(0.041)	(0.037)	(0.031)
Negative Return Dummy		−0.00694	−0.000102	−0.000483		0.000535	0.00750	0.00732
		(0.007)	(0.006)	(0.006)		(0.007)	(0.007)	(0.007)
Extreme Return Dummy		0.0261	0.0191	0.0293		0.00748	0.00246	0.00815
× Negative Return Dummy		(0.054)	(0.054)	(0.056)		(0.046)	(0.045)	(0.044)
Abnormal Volume Quintile		0.00665	−0.00623	−0.0408		0.0147	0.00267	−0.0383
		(0.029)	(0.028)	(0.033)		(0.030)	(0.030)	(0.030)
Abnormal Volatility Quintile		0.0163	0.0195	0.00903		0.0314	0.0318	0.0117
		(0.031)	(0.029)	(0.028)		(0.032)	(0.031)	(0.029)
Up Market		−0.0147*	−0.0130*	−0.0133***		−0.0105	−0.00874	−0.0128**
		(0.008)	(0.008)	(0.005)		(0.008)	(0.008)	(0.005)
VIX (standardized)		−0.0114***	−0.0105***	−0.00889***		−0.0111***	−0.0101**	−0.0111***
		(0.004)	(0.004)	(0.003)		(0.004)	(0.004)	(0.003)
Broker FE	No	No	No	Yes	No	No	No	Yes
Firm FE	No	No	No	Yes	No	No	No	Yes
Observations	63,079	63,079	63,079	63,079	63,079	63,079	63,079	63,079
Adjusted R^2	0.016	0.010	0.018	0.043	0.017	0.011	0.018	0.087

significant but economically small. A one-standard-deviation higher VIX is associated with a 1.1% lower identification rate, which suggests that subpenny price improvement is somewhat less common in volatile markets. We find similar results for the quote midpoint method in column (6).

In columns (3) and (7), we combine the spread variables and controls. The spread indicator and stock price variables have the largest impact on identification. High-priced stocks with large spreads have trades that are less likely to be identified by both algorithms. In columns (4) and (8), we find similar results when we add broker and firm fixed effects.

In summary, the evidence indicates that spreads are the main drivers of identification rates for both algorithms. Predictably, this leads to large variation in identification rates across stocks ranging from 20% to 50% (see Internet Appendix Figure IA.4).

D. Multivariate Analysis of Accuracy Rates

We estimate a similar model to explore variation in conditional accuracy rates. Because the analysis is conditional on identification, the number of trades in each regression is about a third of the number of trades used in the identification regressions. The dependent variable is an indicator variable that takes a value of one if a trade identified as retail is accurately signed as a buy or sell.

Table V summarizes the results. In column (1), we observe that spreads are statistically and economically important determinants of BJZZ accuracy rates. Relative to one-penny spreads, BJZZ accuracy rates drop 15.4 ppt for spreads of 1 to 5 cents and by a very large 40.5 ppt for spreads of 10 cents or more. In column (2), we find that many stock characteristics are related to BJZZ accuracy rates. For example, higher stock prices are associated with lower accuracy rates. More positive returns and extreme positive returns are associated with lower accuracy rates. The interaction with the negative return dummy suggests that the effect of extreme returns is most pronounced for the extreme positive returns. We also observe higher accuracy rates when a stock has abnormally high volume.

Importantly, in column (3), we observe that all of the stock characteristics in column (2) decline in magnitude and mostly lose statistical significance when we introduce the key spread variables. This suggests that spreads are the mechanism that generates the correlations between accuracy rates and stock characteristics. Spreads remain the primary determinants of accuracy rates when we add broker and stock fixed effects in column (4). Overall, we find that the BJZZ algorithm's ability to accurately sign trades declines dramatically for trades that face a spread greater than one penny.

In sharp contrast, the quote midpoint method accuracy rates are more consistent across spread levels. In column (5), we observe that accuracy rates increase modestly as spreads increase, which suggests that price improvement greater than 50% is less likely as spreads increase. Note, however, that accuracy rates are between 94% and 99% for all spread levels (see Figure 2).

Table V

Drivers of the Accuracy Rate: Subpenny Amount (BJZZ) vs. Quote Midpoint Signing

The table presents a linear probability model that examines the drivers that impact the signing conditional accuracy of algorithms using subpenny amount signing (BJZZ) and quote midpoint signing. Our trades are made between December 2021 and June 2022. The sample only includes trades that can be identified as retail trades by the corresponding algorithm. The dependent variable is an indicator that equals one if a trade is signed correctly conditional being identified by one of the algorithms. See Table IV for a description of the independent variables. Standard errors are clustered at the firm and day level and reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Dep Var:	Accuracy Indicator							
Identification Method:	Subpenny PI (exclude [$0.004, $0.006])				Subpenny PI (exclude [40% NBBO, 60% NBBO])			
Signing:	Subpenny Amount (BJZZ)				Quote Midpoint			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spread > $0.01 and Spread ≤ $0.05	−0.154***		−0.159***	−0.137***	0.0352***		0.0329***	0.0244***
	(0.011)		(0.012)	(0.011)	(0.005)		(0.005)	(0.006)
Spread > $0.05 and Spread ≤ $0.10	−0.315***		−0.325***	−0.272***	0.0467***		0.0428***	0.0446***
	(0.016)		(0.018)	(0.023)	(0.006)		(0.007)	(0.008)
Spread > $0.10	−0.405***		−0.420***	−0.370***	0.0298***		0.0233***	0.0321***
	(0.010)		(0.014)	(0.020)	(0.005)		(0.007)	(0.009)
Odd Spread	−0.00707		−0.00786	−0.00735	0.00823**		0.00790**	0.00548*
	(0.009)		(0.009)	(0.009)	(0.003)		(0.003)	(0.003)
Ln(Price)		−0.0521***	0.00724*	−0.0116		0.00500***	0.00268*	0.00427
		(0.006)	(0.004)	(0.018)		(0.001)	(0.002)	(0.010)
Return Quintile		−0.00687**	−0.00198	−0.00162		0.00225	0.00188	0.00178
		(0.003)	(0.003)	(0.003)		(0.001)	(0.001)	(0.001)
Extreme Return Dummy		−0.141*	−0.0902	−0.0737		−0.0733	−0.0810	−0.0659
(=1: absolute return in top 0.5 percentile)		(0.072)	(0.065)	(0.051)		(0.051)	(0.051)	(0.051)
Negative Return Dummy		−0.0321***	−0.00282	−0.00292		0.00433	0.00243	−0.000345
		(0.011)	(0.008)	(0.009)		(0.006)	(0.005)	(0.005)
Extreme Return Dummy		0.177*	0.128	0.124		0.0705	0.0815	0.0734
× Negative Return Dummy		(0.090)	(0.092)	(0.087)		(0.060)	(0.059)	(0.055)

(Continued)

Table V—*Continued*

Dep Var:	Accuracy Indicator							
Identification Method:	Subpenny PI (exclude [$0.004, $0.006])				Subpenny PI (exclude [40% NBBO, 60% NBBO])			
Signing:	Subpenny Amount (BJZZ)				Quote Midpoint			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Abnormal Volume Quintile		0.0906**	0.0542	0.0606**		−0.00663	−0.00605	0.00619
		(0.041)	(0.033)	(0.029)		(0.022)	(0.023)	(0.019)
Abnormal Volatility Quintile		0.0677	0.0589*	0.0307		0.0119	0.0119	0.0122
		(0.044)	(0.035)	(0.038)		(0.018)	(0.019)	(0.019)
Up Market		−0.00584	−0.000374	−0.00159		0.00293	0.00252	−0.00236
		(0.008)	(0.007)	(0.007)		(0.004)	(0.004)	(0.005)
VIX (standardized)		−0.00127	0.00000941	−0.000581		−0.00158	−0.00163	−0.000669
		(0.003)	(0.003)	(0.003)		(0.002)	(0.002)	(0.002)
Broker FE	No	No	No	Yes	No	No	No	Yes
Firm FE	No	No	No	Yes	No	No	No	Yes
Observations	21,810	21,810	21,810	21,810	20,868	20,868	20,868	20,868
Adjusted R^2	0.113	0.031	0.113	0.134	0.004	0.001	0.004	0.121

In column (6), the only statistically significant stock characteristic is the price level, but this effect is considerably reduced when we introduce the spread variables in column (7) and loses statistical significance when we introduce broker and stock fixed effects in column (8).

Predictably, as shown in Internet Appendix Figure IA.5, the BJZZ measure generates large variation in accuracy across stocks (ranging from about 50% to 100%). In contrast, the quote midpoint method materially shrinks the variation across stocks (ranging from about 85% to 100%).

E. Robustness Checks

We supplement our main analysis with several robustness checks to ensure that the results regarding identification rates and accuracy rates are reasonable. To do so, we run auxiliary experiments with different trade types or targeting specific stocks. Table VI uses the column (1) regression model of identification and accuracy rates from Tables IV and V, respectively, but adds an indicator variable for the trade or stock of interest.

To address the concern that our small trades yield identification and accuracy rates that are different than those for larger trades, we execute 2,292 trades of about $1,000 in 130 randomly selected stocks that are in our main sample. The experiment compares the identification and accuracy rates of small and large trades in the same stock on the same day. To limit the cost of this experiment, we do not trade in all 128 stocks each day. As for our main sample, we choose a share size that yields a trade value closest to $1,000. In Table VI, Panel A, we find that both algorithms have similar identification rates for these larger trades (columns (1) and (3)). BJZZ accuracy rates are lower for these trades, but the economic magnitude is small (4.9 ppt lower on a baseline accuracy rate of 72%). The quote midpoint method yields a smaller reduction in accuracy rates; the reduction is economically small and no longer statistically significant.[13]

Bartlett (2021) shows that there is a sharp increase in de minimis price improvement for trades of 100 shares relative to odd-lot trades of less than 100 shares. He attributes the difference to regulations that require market centers to post the percentage of trades that receive price improvement for trades of 100 shares or more. To see if this pattern materially affects our inferences, we run an experiment in which we execute 338 trades of 100 shares alongside the $100 trades in the main sample. In Table VI, Panel B, we observe higher identification rates for both algorithms, which is consistent with the increased prevalence of de minimis subpenny price improvement for 100-share trades. Accuracy rates are unaffected.

[13] Somewhat surprisingly, this is because price improvement greater than 50% is somewhat more likely for these larger trades. For buys, 15.2% of both large and small trades experience price improvement over 50%. For sells, 15.8% of large trades and 11.0% of small trades experience price improvement over 50%. The quote midpoint method excludes a larger region of trades around the quote midpoint and thus reduces the inaccuracy that occurs in signing trades with greater than 50% price improvement.

Table VI

Robustness Checks

The table examines a number of multivariate regressions of identification and conditional accuracy rates with an indicator variable for the stock or trade characteristic of interest. The analyses are based on our trades made between December 2021 and June 2022. Columns (1) and (2) report results for the algorithm using subpenny amount signing (BJZZ), and columns (3) and (4) report results for the algorithm using quote midpoint signing. Panel A includes $1,000 trades and corresponding $100 trades for the same stock on the same day. Larger Trade is an indicator for trades of about $1,000. Panel B includes 100-share trades and corresponding $100 trades for the same stock on the same day. 100-share Trade is an indicator for trades with 100 shares. Panel C includes all top mover trades with trades from our main sample. Each day, four stocks on the Robinhood Top Mover list with share prices greater than $1 are identified. Top Mover Day is an indicator for the trading day immediately following the 9 am observation of the large overnight move. Post Top Mover Day is an indicator for the four days immediately following the top mover day. Panel D includes trades of retail darling stocks with trades from our main sample. Retail Darling is an indicator for stocks that are among the 30 most popular holdings on Robinhood. Panel E includes trades of megacap stocks with trades from our main sample. Mega Cap is an indicator for megacap stocks. Control variables include trade characteristics: indicators for different bid-ask spread regions (($0.01, $0.05], ($0.05, $0.10], and greater than $0.10) and an indicator for spreads that are odd number cents. Standard errors are clustered at the firm and day level and reported in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Identification Method:	Subpenny PI (Exclude [$0.004, $0.006])		Subpenny PI (Exclude [40% NBBO, 60% NBBO])	
Signing:	Subpenny Amount (BJZZ)		Quote Midpoint	
Dep Var:	Identification Indicator	Accuracy Indicator	Identification Indicator	Accuracy Indicator
	(1)	(2)	(3)	(4)
	Panel A: Larger trade			
Larger Trade	0.0182	−0.0492**	0.0293	−0.0152
	(0.016)	(0.024)	(0.018)	(0.009)
Observations	4,663	1,702	4,663	1,630
Adjusted R^2	0.009	0.110	0.010	0.009
	Panel B: 100-share trade			
100-Share Trade	0.148***	0.0179	0.0997***	−0.00269
	(0.014)	(0.049)	(0.012)	(0.003)
Observations	673	211	673	205
Adjusted R^2	0.021	0.082	0.010	−0.003
	Panel C: Top mover			
Top Mover Day	0.00329	−0.0284	0.00972	−0.0348**
	(0.015)	(0.019)	(0.015)	(0.014)
Post Top Mover Day	0.0241**	−0.00155	0.0318***	−0.0131*
	(0.011)	(0.009)	(0.011)	(0.007)
Observations	69,608	24,340	69,608	23,379
Adjusted R^2	0.016	0.114	0.016	0.005

(*Continued*)

Table VI—*Continued*

Identification Method:	Subpenny PI (Exclude [$0.004, $0.006])		Subpenny PI (Exclude [40% NBBO, 60% NBBO])	
Signing:	Subpenny Amount (BJZZ)		Quote Midpoint	
Dep Var:	Identification Indicator	Accuracy Indicator	Identification Indicator	Accuracy Indicator
	(1)	(2)	(3)	(4)
	Panel D: Retail darling			
Retail Darling	0.0224	−0.000108	0.0342**	−0.0370***
	(0.017)	(0.023)	(0.015)	(0.009)
Observations	67,007	23,355	67,007	22,435
Adjusted R^2	0.015	0.114	0.016	0.006
	Panel E: Mega cap			
Mega Cap	−0.0214	−0.0277	0.00278	−0.0218
	(0.021)	(0.018)	(0.016)	(0.018)
Observations	66,680	23,125	66,680	22,232
Adjusted R^2	0.016	0.114	0.017	0.005

Next, several studies document that retail investors tend to be net buyers of stocks on days with extreme returns (Barber and Odean (2008), Barber et al. (2022)). To assess the performance of the algorithm on those days, we run an experiment in which we place 5,509 trades in stocks that appear on the Robinhood Top Mover list each day before the market opens (i.e., the stocks have larger overnight moves) and have share prices greater than $1. By luck, only 6.6% of these overnight return moves are negative. We compare the performance of the algorithm in these stocks relative to our main sample on the "Top Mover Day" (the trading day immediately following the 9 a.m. observation of the large overnight move). We also analyze the performance "Post Top Mover Day" (the four days immediately following the top mover day). This regression (and those in the remaining panels) contains all observations from our main sample because the experiment investigates stock (rather than trade) characteristics. In Panel C, we find that identification rates are not materially affected on these top mover days but increase modestly in the days following (columns (1) and (3)). Accuracy rates are somewhat lower on the top mover days. However, the economic magnitude of the effect is small (e.g., −3.5 ppt on a baseline accuracy rate of 95% for the quote midpoint algorithm). The small reduction in accuracy can be traced to the higher frequency of 50% or greater price improvement for sell trades, perhaps because large price moves attract buying activity, so sellers receive good price execution.

We next analyze "retail darlings." We identify the stocks in our main sample that are among the 30 most popular holdings on Robinhood as of May 2022, adding Apple and Tesla, which are also in the Robinhood top 30. Panel D shows the identification and accuracy rates for these retail darlings. There is some evidence of higher identification rates, but the effect is economically small. Among retail darlings, the quote midpoint method has somewhat lower accuracy rates but still dominates the BJZZ method given the high baseline accuracy rate of 95% for the quote midpoint method.

Finally, we perform a similar analysis for megacap stocks since most trading volume is concentrated in large cap stocks. Specifically, we execute targeted trades in six megacap stocks (i.e., AAPL, BAC, GOOG, NVDA, V, XOM) with market caps greater than $700 billion at the time of sample selection. By comparison, the largest stock in our main sample is Goldman Sachs with a market cap of $129 billion. Panel E shows that identification rates and accuracy rates are similar for these megacap stocks.

III. Discussion

A. False Positives

In the analysis above, we show that false negative identification is common; indeed, 65% of our trades are not identified as retail. However, our data set cannot measure the frequency of false positives (institutional trades that execute off exchanges at subpennies). Our discussions with SEC staff, FINRA staff, and traders at financial institutions suggest that institutional orders are not typically sold to wholesalers. Institutions do trade off exchanges through crossing networks and alternative trading systems (ATS). However, most of those we spoke with thought that institutional trades are more likely to execute in round pennies. Quote midpoints are an exception: BJZZ (p. 2255) note that "…Reg NMS has been interpreted to allow executions at the midpoint between the best bid and best offer. As a result, institutions are heavy users of crossing networks and midpoint peg orders that generate transactions at this midpoint price."

BJZZ (p. 2251) argue that "…trades at or near a half-penny are likely to be from institutions and are not assigned to the retail category." If true, including trades near the quote midpoint would increase the frequency of false positives. Thus, we also exclude trades that execute within 40% and 60% of NBBO when measuring retail trade. This exclusion is justified for measuring retail order imbalance because signing accuracy drops near the quote midpoint. However, whether excluding trades near the quote midpoint degrades or improves measures of the intensity of retail trading remains an open question.

To shed some light on these questions, we compare the distribution of our trades with all off-exchange trades (marked with code "D" on TAQ). We find that our trades are less likely to occur at a round penny than other off-exchange trades (39.30% of our trades occur at a round penny vs. 50.58% of all off-exchange trades), perhaps because off-exchange institutional trades are more likely to execute at a round penny. After we exclude trades at round pennies, the distribution of our trades and all off-exchange trades are similar (see Internet Appendix Figure IA.6). Our trades are less likely to execute at $0.005, which is often the quote midpoint and might contain an outsized fraction of institutional trades as discussed above. We also observe somewhat fewer observations in our trades in the low and high subpenny digits, perhaps because subpenny price improvement is more common for round lots, as we document in robustness checks, given that our trades are odd lots. We do not view this evidence as definitive. In summary, whether using the subpenny digit in off-exchange trades to identify retail trades leads to a large rate of false positive errors is an open question.

B. Excluding Trades Near the Quote Midpoint

We show that signing accuracy is close to 50% around the quote midpoint (Figure 3, Panel C) which suggests that trades between 40% and 60% of the quote midpoint be excluded to reduce signing errors when measuring retail order imbalance. When a researcher excludes trades from the calculation of order imbalance, one implicitly assumes that order imbalance for the included trades equals order imbalance for the excluded trades. In Internet Appendix Figure IA.8, we show this is a reasonable assumption across a wide range of order imbalance using data from our trading experiment.

However, at extreme levels of order imbalance, excluding trades near the quote midpoint inflates measures of order imbalance. For example, on days with heavy buying, sell orders are more likely to execute near the quote midpoint, and vice versa. However, in Internet Appendix Section IV and Figure IA.8, we show that this inflation bias is relatively small.[14] In the same section, we offer an alternative order imbalance measure that assumes zero order imbalance for the excluded trades and thus "shrinks" order imbalance estimates toward zero. Although the alternative measure has greater absolute error, the two measures are nearly perfectly correlated. Thus, they will rank stocks in similar ways and generate similar test statistics in regression analyses. However, scholars interested in the magnitude of order imbalance might use both measures to provide reasonable bounds on order imbalance.

IV. Conclusions

BJZZ develop an algorithm that identifies retail trades that execute at subpenny amounts on off-exchange venues. The method has quickly become common in the academic literature. We examine the effectiveness of the BJZZ subpenny digit algorithm and introduce a quote midpoint method to identify and sign retail trades using a self-generated set of over 85,000 retail trades. With respect to identification, both algorithms use subpenny price execution to identify trades. The BJZZ algorithm excludes trades that execute at subpennies between \$0.004 and \$0.006; the quote midpoint method excludes trades that execute between 40% and 60% of the quote midpoint, relative to NBBO. Both methods correctly identify about a third of our trades as retail trades. However,

[14] For example, when estimated retail order imbalance is −0.73 (86.5% sells), the bias is -0.05 (i.e., true order imbalance is -0.66 or 83% sells).

because subpenny price improvement is less likely as spreads increase, identification rates vary systematically across stocks. The variation mechanically generated by spreads must be addressed to avoid false discoveries regarding the intensity of retail trading.

In terms of accuracy of buy or sell signing, the BJZZ subpenny digit method incorrectly signs 28% of identified trades (i.e., has an accuracy rate of 72%). As with identification rates, accuracy rates decrease as spreads increase, vary systematically across stocks, and yield uninformative measures of order imbalance for 30% of the stocks in our sample. Fortunately, our preferred method—using the quote midpoint to sign trades while excluding trades that executed within 40% to 60% of the NBBO—improves accuracy to over 95% and greatly reduces variation in accuracy rates across stocks and spread levels.

Initial submission: September 28, 2022; Accepted: June 21, 2023
Editors: Antoinette Schoar, Urban Jermann, Leonid Kogan, Jonathan Lewellen, and Thomas Philippon

REFERENCES

Barber, Brad, Xing Huang, Terrance Odean, & Christopher Schwarz, 2022, Attention induced trading and returns: Evidence from Robinhood users, *Journal of Finance* 77, 3141–3190.

Barber, Brad, Shengle Lin, & Terrance Odean, 2023, Resolving a paradox: Retail trades positively predict returns but are not profitable, *Journal of Financial and Quantitative Analysis*, forthcoming.

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

Bartlett, Robert, 2021, Modernizing odd lot trading, *Columbia Business Law Review* 2021, 520–568.

Battalio, Robert, Robert Jennings, Mehmet Saglam, and Jun Wu, 2022, Identifying market maker trades as "retail" from TAQ: No shortage of false negatives and false positives, Working paper, University of Notre Dame.

Blankespoor, Elizabeth, Ed Dehaan, John Wertz, and Christina Zhu, 2019, Why do individual investors disregard accounting information? The roles of information awareness and acquisition costs, *Journal of Accounting Research* 57, 53–84.

Boehmer, Ekkehart, Charles Jones, Xiaoyan Zhang, and Xinran Zhang, 2021, Tracking retail investor activity, *Journal of Finance* 76, 2249–2305.

Bonsall, Samuel, Jeremiah Green, and Karl Muller, 2020, Market uncertainty and the importance of media coverage at earnings announcements, *Journal of Accounting and Economics* 69, 101264.

Bradley, Daniel, Russell Jame, and Jared Williams, 2022, Non-deal roadshows, informed trading, and analyst conflicts of interest, *Journal of Finance* 77, 265–315.

Bushee, Brian, Matthew Cedergren, and Jeremy Michels, 2020, Does the media help or hurt retail investors during the IPO quiet period? *Journal of Accounting and Economics* 69, 101261.

Farrell, Michael, Clifton Green, Russell Jame, and Stanimir Markov, 2022, The democratization of investment research and the informativeness of retail investor trading, *Journal of Financial Economics* 145, 616–641.

Guest, Nicholas, 2021, The information role of the media in earnings news, *Journal of Accounting Research* 59, 1021–1076.

Israeli, Doron, Ron Kasznik, and Suhas Sridharan, 2022, Unexpected distractions and investor attention to corporate announcements, *Review of Accounting Studies* 27, 477–518.

Lee, Charles, and Mark Ready, 1991, Inferring investor behavior from intraday data, *Journal of Finance* 46, 733–746.

Levy, Bradford, 2022, Price improvement and payment for order flow: Evidence from a randomized controlled trial, Working paper, University of Chicago.

Schwarz, Christopher, Brad Barber, Xing Huang, Philippe Jorion, and Terrance Odean, 2023, The "actual retail price" of equity trades, Journal of Finance (forthcoming).

Securities and Exchange Commission, 2005, Regulation NMS, SEC Release No. 34–51808.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication Code.