# Trade Size and the Cross-Section of Stock Returns

**David A. Lesmond***

November 18, 2023

[1]David Lesmond (dlesmond@tulane.edu) is affiliated with the Freeman School of Business. Please do not cite, circulate or quote this paper at this time

# Trade Size and the Cross-Section of Stock Returns

**Abstract**

This paper finds that retail trading can be identified by medium sized trades between 500 and 999 shares. In particular, medium trade sizes are priced in future monthly NYSE/Amex returns. The results indicate that firms with low levels of medium sized trades outperform those with high levels and earn approximately 50 basis points (bps) per month over the five-factor Fama-French model. These significant pricing results persist regardless of pre and post 2001 decimalization in stock quotes or after controlling for independent characteristics (Green, Hand, and Zhang, 2017; Lewellen, 2015) shown to be related to future returns. The results add further empirical evidence on the identification and pricing of retail trading.

*Keywords:* Market Microstructure; Trade Size; Intra-Day Trades, Pricing

# 1 Introduction

A considerable body of academic research investigates the identification and pricing implications of individual trading. One line of research utilizes proprietary data to identify a subset of retail investors (Odean, 1998; Kaniel, Saar, and Titman, 2008; Kelley and Tetlock, 2013). Unfortunately, these proprietary databases are separate and distinct across retail brokers and time periods reducing the power of asset pricing tests and its applications. A second line of research relies on the assumption that dollar trading volume is a viable retail trading proxy (e.g. Lee (1992); Hvidkjaer (2006, 2008)). To date, the only empirical evidence validating this claim is provided by Lee and Radhakrishna (2000) who focus on a very short time interval to test this assumption and Hvidkjaer (2008) who confirm dollar trades as a viable proxy for retail trading over a longer time period. However the advent of decimalization reduces the effectiveness of dollar volume as a retail trading proxy limiting these tests and pricing implications to the pre-decimalization period. Finally, Boehmer, Jones, Zhang, and Zhang (2021) propose and develop a methodology to infer retail purchases and sales from publicly available data in the post decimalization period by identifying trades with sub-penny prices that are reported to FINRA's Trade Reporting Facility (TRF) off-exchange.

This paper finds that medium sized trades, namely between 500 and 999 shares, are highly associated with the commonly used proxies for retail trading spanning dollar trades and off-exchange trading platforms. I further find that these retail trades are negatively and significantly associated with future returns asross the pre and post-decimalization periods. In other words, stocks favored by retail investors tend to experience under-performance in association with future monthly returns. Finally, the pricing of medium, retail trades is robust to alternative explanations spanning 22 known anomalies indicating a separate and distinct pricing anomaly.

Cready (1988); Cready and Mynatt (1991); Barclay and Warner (1993) and Battalio and Mendenhall (2005) show that different types of investors[1] can be identified by trade size given by the number of trades in a particular category. However, many studies rely on dollar trade value

[1] There is long-standing empirical evidence of systematic trading behavior among various investor groups. For instance, small and large investors respond differently to exogenous information events such as earnings releases (Lee, 1992), seasoned equity offerings (Huh and Subrahmanyam, 2005), and analyst recommendations (Malmendier and Shanthikumar, 2014).

to examine pricing first exemplified in Lee (1992). While this is an effective means of identifying small trades, its usefulness is limited to the breakdown in trade size due to the decimalization in stock quotes (pre-decimalization) beginning in 2001. Indeed papers [see for example, (Campbell, Ramadorai, and Schwartz, 2009)] cease their analysis at 2000 reflecting this concern.

I examine trade size clusters, rather than a continuous measure of trade size, because Barclay and Warner (1993); Chakravarty (2001) document that specific medium-sized trades have the most cumulative price impact which they interpret as having more value relevant information. This is further exemplified by Alexander and Peterson (2007) who note investor clienteles cluster at 100, 500 and 1000-shares per trade. I break down the medium-sized trades using trade size clusters between 500 to 999, 1000 to 4999, and 5000 to 9999 shares to split out medium sized trades, while keeping 100 to 499 share trades as signifying small trades. Trades larger than 10000 shares are deemed large trades.

I first document that medium sized trades (between 500 and 999 shares) are significantly associated with the commonly used $10,000 limit where I show significant association with the proportion of retail trades. The relation is robust and positive indicating that higher levels of medium trades are associated with a greater proportion of retail trading activity. I then examine the association of trade size and off-exchange buy and sell order imbalance. I find that medium trades are again priced with future weekly returns indicating that retail traders use medium sized trades (in addition to small trades). Hence, medium sized trades are shown to be associated with retail trading across the pre and post-decimalization period. However, I also document that over monthly trading horizons this pricing effect is now negatively and severely weakened.

I find that medium-sized trades produce a negative five-factor Fama-French alpha of approximately 60 basis points (bps) per month. The significance of alpha is maintained across both the pre and post-decimalization[2] periods, but falls from -61 bps to -37 bps per month, respectively.

[2] The NYSE Fact book reports statistics showing average trade sizes falling dramatically after stock decimalization. The average trade size in 1999 for NYSE-listed firms was 1,205 shares per trade. After decimalization in 2004, the average trade size was significantly reduced to just over 390 shares per trade. In 2010, the average trade size had dwindled to 220 shares per trade and in 2019 the average trade size was approximately 140 shares per trade. The post-decimalization period also affected volume measures as investors significantly increased trade volume in reflection of reduced trading costs.

No other trade size classifications are consistently priced in future monthly returns. I relate order imbalance and these medium-sized trades and find that trade flows favors low medium-sized trades for the pre-decimalization period earning 0.44% per month. This behavior shifts significantly in the post-decimalization period where the trade imbalance now favors high medium-sized trades but loses -0.41% per month. I interpret these results to show that medium, retail trades are largely uninformed at monthly trading horizons, and display evidence of stealth trading (Barclay and Warner, 1993).

These results are robust to various controls for the pricing of future monthly returns. Using both Green et al. (2017) and Lewellen (2015) as control variables, I find that in sort tests separate controls for the pricing of medium trades do not obviate its impact. The results show continued negative and significant pricing with future monthly returns regardless of the control employed. The results are dependent on small firms and appear to be most prominent for the long side of the trade. These results extend to Fama and MacBeth (1973) regression tests that show a negative and robust association of medium trades to future monthly returns over the total sample period and various sub-periods.

Our results extend the recent literature on retail investors. In particular, the evidence shows that medium sized trades (Lee, 1992; Hvidkjaer, 2006, 2008) and off-exchange trade clearing (Boehmer et al., 2021) may now find a common thread to retail trading. Building upon those earlier findings, we identify a proxy of retail trading for a longer time period and use it to study the asset pricing implications of speculative retail clienteles. Our results add to the growing evidence on the importance of retail investors in the return-generating process (e.g., Kumar and Lee (2006), Barber, Odean, and Zhu (2009), Dorn, Huberman, and Sengmueller (2008), Hvidkjaer (2008), and Kaniel, Saar, and Titman (2008)). In broader terms, our results highlight the usefulness of a habitat-based approach for studying asset prices.

This study is important for the following reasons. We extend the line of research into trade size, but in a very different mode. By obviating the necessity of identifying trade direction and focusing on unidirectional trade size, we show that an easily estimable trade size portfolio can enhance the profitability of trade based strategies. Rather than focusing on trade imbalances that are institutionally based (Kaniel, Liu, Saar, and Titman, 2012) or utilizing intraday dollar-volume based small trades (Lee, 1992), we would contend that medium trade cluster portfolios are related to a vast array of anomalies (Hou, Xue, and Zhang, 2015).

The paper is organized as follows. Section 2 outlines the estimation of trade size portfolios and the various control variables, and Section 3 outlines the summary statistics in the pre and post-decimalization periods. Section 4 outlines trade size determinants and Section 5 presents decomposition of medium trades by dollar volume and off-exchange trading. Section 6 shows the single sort five-factor Fama-French alphas, and Section 7 presents the pre and post-decimalization sort tests. Section 8 illustrates the relation between trade size and order imbalance. Section 9 shows double sort tests with various control variables, and Section 10 shows the Fama-MacBeth regression tests. Finally, Section 11 shows the relation between cumulative abnormal returns around earnings announcements and trade size. Section 12 concludes.

## 2 Data

The sample includes all ordinary common stocks listed on the NYSE and American Stock Exchange (Amex)[3] over the period January 1980 through December 2022. I focus on both the Institute for the Study of Security Markets (ISSM) and the Trade and Quote (TAQ) data to obtain trade size and the national best bid-offer (NBBO) bid-ask spread relatable to each trade. I augment ISSM and TAQ with Francis Emory Fitch intraday trade data [4] across both NYSE and Amex listed firms.

The Fitch database includes all NYSE/Amex trades from 1980 to 1982, the ISSM data set includes all NYSE/Amex trades from 1983 to 1992, while TAQ covers 1993 to 2022 for all exchanges. Trades with irregular terms are excluded and trades are run through a simple price-based error filter to exclude likely erroneous prices.

[3]Ellis K. and M. (2000) show that the use of trade classification rules such as Lee and Ready (1991) in Nasdaq data introduces biases in classifying large trades and trades initiated during high-volume periods, especially for trades executed inside the spread.

[4]Fitch data do not contain bid-ask spreads.

Using Barclay and Warner (1993) as a basis, I initially calculate the percentage of trades among small (100 to 499 shares), medium (500 to 9,999 shares), and large trade (great than 10,000 shares) categories. However, Alexander and Peterson (2007) find that trades cluster at 500, 1,000, and 5,000 shares. Using these trade size clusters as a boundary, I further break down trades between 500 to 9,999 share into 500 and 999 shares, 1,000 to 4,999, and 5,000 to 9,999 share categories. I count the trades within each category and then divide this sum by the total number of trades each day. The daily percentage is then averaged over the month.

I also calculate trade direction using the Lee and Ready (1991) algorithm by matching the NBBO bid-ask spread to each trade for both ISSM and TAQ to determine buys and sells. For the Fitch data, I only use the tick test to classify trade as buys or sells. The daily percentages are calculated as (buys-sells)/(buys+sells) trades each day and then averaged over the month.

Firm specific characteristics are estimated using the Green et al. (2017) methodology [5] and Lewellen (2015). I combine the characteristics from these two papers, but focus primarily on the 12 firm characteristics that Green et al. (2017) show are independent in asset pricing. I then add the characteristics contained in Lewellen (2015), as additional firm controls that have been shown to be related to future security returns. Return data and unsigned end of day share volume data are from the Center for Research in Security Prices (CRSP) files.

Finally, I delete all stocks with a price less than $5 during the prior month that mitigates any microstructure issues or regulatory concerns that may impede investability by institutions. The results are robust to alternative price filters.

I implement an 1/0/1 trading strategy that estimates trade size in month t and all the remaining control variables in month t-1. I do not exclude the month between the estimation and performance period.

[5] I thank the authors for providing the computer code to estimate each of the characteristics used in the paper.

## 3 Summary Statistics

Panels A and of Table 1 reports the descriptive statistics of the sample where Panel A presents the pre-decimalization period, 1980m1 to 2000m12, and Panel B presents the post-decimalization period, 2001m1 to 2021m12. Within each panel, I decompose average trade size into separate categories, where trades are grouped into portfolios using established (Barclay and Warner, 1993) trade size clusters between 100 to 499 shares, 500 to 999 shares, 1000 to 4999 shares, 5000 to 9999 shares, and finally $\geq$ 10000 shares.

Panel A of Table 1 reports an average of 78.15 daily trades over the month with a standard deviation of 64.25. Over the month, there are a minimum of 18.32 and a maximum of 327.15 trades per day over the month. Turning to trade size we see that as found in Barclay and Warner (1993) that small trades dominate the trading pattern. 54.72% of trade are less than 4,999 shares, whereas 18.51% and 21.8% of trades are between 500 and 999 or 1000 and 4999 shares, respectively. The remaining trade categories show 2.61% and 1.9% for share trades greater than 5000 and 10000 shares, respectively. The remaining statistics show that within each trade size category the distribution is tightly centered around the mean as evidenced by the low standard deviations.

Turing to Panel B of Table 1, it is clearly evident that the decimalization in stock quotes led to an explosion in trades. During this period, there are 6101.80 daily trades over the month across all NYSE/Amex stocks. This is an almost ten fold increase in trades from the pre-decimalization period. Again, small trades dominate the trading behavior with over 71% of all trades occurring between 100 and 499 shares. The increase in small trades is at the expense of larger trades where we see only 6.5% and 5.7% of trades between 500 to 999 and 1000 to 4999 shares, respectively. There are less than 1% of trades for more than 5000 shares. Also evident in Panel B is a far greater standard deviation in trades across all of the trade size categories.

# 4 Trade Size Determinant Tests

I now empirically investigate what drives each trade size portfolio. Specifically, I examine how trade size portfolios are related to past order flow and past returns. I adopt the Fama and MacBeth (1973) two-stage estimation with zero lags to provide maximum flexibility and focus on cross-sectional patterns. I use the past one month return to explain the dependency of trade size on the prior month's return and the 11-month momentum returns to examine if there is either persistence or a contrarian behavior in prior returns. I also use Brennan and Subrahmanyam (1998)[6] and Boehmer et al. (2021) as a basis for the explanatory variables that include the number of institutions, the percentage of institutional[7] holdings, turnover (variable cost estimate), the percentage of zero returns (the fixed cost element), return volatility, book-to-market, and firm size. These variables are all log scaled and lagged by one month.

Table 2 presents the results, with regressions covering the 100 to 499, 500 to 999, 1,000 to 4,999, 5,000 to 9,999, and ≥10,000 share categories. As is shown, there is strong, positive, and significant persistence across the share categories. The coefficients range from 0.8115 to 0.6574 for the 100 to 499 and ≥ 10,000 share categories, respectively. The persistence in trade size is consistent with Chordia and Subrahmanyam (2004) who report a similar finding for order imbalance.

The coefficients for the past month momentum ranges from -0.0073 to -0.0031 for share trades from 100 to 499 and ≥ 10,000, respectively, and past 11-month momentum that ranges from -0.0012 to zero for share trades from 100 to 499 and ≥ 10,000, respectively¿ The negative sign for all medium trades indicates a contrarian behavior with past returns. This is opposite to that obtained for the 100 to 499 share category that shows distinctive continuation.

Turning to the control variables, I first focus attention on the institutional aspect by noting that

[6] Brennan and Subrahmanyam (1998) uses intraday estimates of the variable and fixed costs elements. In place of these intraday estimates, I substitute a low frequency variable cost turnover measure and I use the fixed cost zero return measure Lesmond, Ogden, and Trzcinka (1999).

[7] Daniel, Grinblatt, Titman, and Wermers (1997) show that stocks held by mutual funds outperform a variety of benchmarks indicating informed behavior, while Grinblatt and Titman (1989) and Daniel et al. (1997) find that mutual finds do display some selectivity ability again indicating informed behavior. More broadly, Gompers and Metrick (2001) show that the fraction of a firm's shares held by all institutions predicts returns cross-sectionally after controlling for other firm characteristics. However, Lewellen (2011) finds little evidence of abnormal performance for mutual funds.

only the 500 to 999 share trade category shows negative and significant coefficients for both the level of institutional holdings and the number of owners. Decreased levels of institutional interest is consistent with increased levels of retail trading. It should be noted that the positive levels of institutional interest are only evident for share trades greater than 5,000 shares.

Overall the control variables indicate that small trade (100 to 499 shares) investors tend to buy more aggressively in smaller firms, value firms, and firms with lower turnover and lower return volatility, while 500 to 999 share investors tend to concentrate in smaller firm, growth firms, and firms with higher return volatility. These are distinctly reversed for larger trade sizes. All coefficients are highly significant. The average adjusted $R^2$ is approximately 12% from the first stage cross sectional estimation.

The results in Table 2 reveal three important drivers affecting the the size of share trades. First, institutional holdings and firm size are distinctive for different trade size categories indicating possible retail trading for traders using 100 to 499 and 500 to 999 share trades. The second is its own lag, which indicates that the trade size measures are persistent. The third are past returns, show a mixed result of both contrarian and momentum patterns, with the contrarian pattern prevailing for medium to large trades and a continuance pattern for small trades.

## 5 Retail Trading Decomposition

I now breakdown two known proxies for retail trading, namely dollar trade size Lee (1992) and off-exchange trade reporting facility trades Boehmer et al. (2021). These tests span the period 1980 to 2000 for dollar trade size and from 2010 to 2015 for the off-exchange retail trades. The first test uses a Fama and MacBeth (1973) procedure to test which trade size categories are associated with dollar trade volume. The commonly used dollar limit is $10,000. Specifically, the retail investors' trading intensity or retail trading proportion (RTP) for a given stock in a certain month is defined as the total retail trading volume divided by the total trading volume in the market (Brandt, Brav, Graham, and Kumar, 2010), and classify them as retail trade.

The second test uses a sort test on a restricted sample of only those trades that are off-exchange trading by retail traders (Boehmer et al., 2021) and examine the pricing details when splitting by trade size. For this test I concentrate on buys minus sells or order imbalance within each trade size category.

Panel A of Table 3 presents the results for the proportion of retail trades on trade size. As shown in Panel A, surprisingly the 100 to 499 share category is negatively associated with the proportion of retail trades showing a coefficient of -0.1371. However, the medium 500 to 999 and 1,000 to 4,999 trade sizes show a positive and significant association with the retail proportion. The coefficients range from 0.3137 (t-statistic of 4.95) to 0.2044 (t-statistic of 2.60) indicating that larger medium trade size volumes are associated with higher levels of retail trading. As expected, much larger trade sizes are negatively associated with retail trading. As shown, the coefficients for trade sizes greater then 5,000 range from -1.0560 to -1.7554, both of with are robustly significant.

Panel B of Table 3 splits the tests into weekly and monthly horizons for returns. Boehmer et al. (2021) uses weekly horizons for their tests. For brevity, I only report the trade sizes that demonstrate significant pricing with future returns. As shown in Panel B, positive pricing (Buys - Sells) is noted across the small and medium trade sizes indicating that retail traders employ small and medium trade sizes in implement their positions. As shown, the small trade alpha is 0.133% monthly and medium trade alphas are 0.099% and 0.065%. The largest magnitudes are centered in the small and 500 to 999 share trade sizes and buys ear a significantly higher alpha than do sells across each trade size.

However, if turn to a monthly horizon, we clearly see that the pricing of retail trading is now negative and it is concentrated in the 100 to 499 (small) and 500 to 999 (medium) trade sizes. The alphas are -0.300% to 0.234% per month for the small and medium size trades, respectively. For these trade size, the sells earn more than the buys, with the sell alpha being robustly significant.

I show in these tests that medium trades, namely those between 500 and 999 are associated with the commonly used proxies of retail trading. Hence, I conclude that medium trades are a viable proxy for retail trading and most importantly this proxy is robust to the pre-decimalization and post-decimalization periods. Prior research (Hvidkjaer, 2006, 2008; Barber, Odean, and Zhu, 2009) shows that commonly used proxies for retail trade specifically rely on the pre-decimaliztion period due to identification issues. These results indicate that medium trades, specifically those in the 500 to 999 range, are adequate to identify retail trades across the pre and post decimalization periods.

## 6 Initial Trade Size Predictive Sort Tests

Table 4 presents the outcomes of the initial test employing decile sorts. Value-weighted alphas have been calculated using a five-factor Fama-French model, with the prior month's firm size serving as the weighting factor. The analysis covers January 1980 (1980m1) to December 2022 (2022m12). Trades have been categorized into distinct trade size groupings, with the 'Low' and 'High' portfolios comprising stocks exhibiting the lowest and highest concentration of share trades, respectively, within each trade size cluster. The results are reported based on trade size categories that encompass the following ranges: 100 to 499, 500 to 999, 1,000 to 4,999, 5,000 to 9,999 shares, and those exceeding 10,000 shares.

Table 4 presents distinct results for each trade size category. First, examining the High-Low quintile sort, shows no significant pricing difference across all trades, indicating an alpha of -0.237% (with a t-statistic of 1.59). This suggests that trade size, on its own, does not appear to be a significant factor in predicting future monthly returns. This lack of significance extends to the 100 to 499 share category, which shows an alpha of 0.251% (with a t-statistic of 1.55). However, the trend in alphas across the 100 to 499 share deciles clearly shows that more informed traders concentrating in small trades earn 0.330% per month.

However, the significance of trade size in pricing becomes evident in the medium-size trade categories, specifically those ranging from 500 to 9,999 shares. Within this range, the 500 to 999 share category demonstrates an alpha of -0.631% per month with a robust t-statistic of 4.8, equating to an annualized abnormal return of 7.60%. As noted by Harvey, Liu, and Zhu (2015), verifiable anomalies should exhibit a robust t-statistic exceeding three, and these results surpass that threshold.

The viability of this pricing lies in the stocks that are avoided by traders using 500 to 999 shares. The low trade size alpha is 0.459% per month (significant at the 1% level), while the high trade size alpha yields -0.171% per month (significant at the 10% level). It is worth mentioning that the trend in alphas across the low to high deciles is nearly monotonic, suggesting that sorting based on 500 to 999 share trades leads to progressively diminishing alphas. The negative sign in the high-low alpha appears to suggest retail trading, namely highly concentrated 500 to 999 share trades earn significantly lower alphas.

Negative and statistically significant alphas are also observed in the 1,000 to 4,999 and 5,000 to 9,999 share trade categories, with high-low alpha of -0.391% per month and -0.415% per month, respectively. As found for the 500 to 999 share category, the stocks with low trade levels yield high alphas. But the monotonicity in alphas across the trade size rankings is not evident for the 1,000 to 4,999 and 5,000 to 9,999 share categories. It is also important to note that the robustness of these findings is notably diminished when compared to the results for trades involving 500 to 999 shares. Notably, the trend across the deciles is non-monotonic, suggesting that sorts using these trade size categories exhibit a weaker association with future returns.

Finally, stocks sorted based on 10,000 share trades do not show significant pricing, displaying an insignificant alpha of -0.242% per month, although again the most pronounced pricing is shown in the lowest level of 10,000 share trades.

It appears that the significant alphas demonstrated by the medium trades is offset by the small trades (100 to 499 shares). The results clearly indicate that stocks with high concentrations of 500 to 999 shares earn a significantly lower returns than those with low concentrations, consistent with the presence of retail trading.

## 7 Pre-Post Decimalization Sort Tests

It is well known that the decimalization of stock prices significantly affects the identification and pricing ability of retail trading (Hvidkjaer, 2008), as well as the effectiveness of pricing anomalies (Green et al., 2017). I now split the sample into two periods demarcated by the decimalization in stock quotes that occurred in 2001. The pre-decimalization period is from 1980m1 to 2000m12 and the post decimalization period is from 2001m1 to 2022m12. This roughly splits the sample into two equal time periods. I again present five-factor Fama-French alphas using decile sorts across five separate trade size groupings.

As shown in Table 5, across all trades, there is no significance in the pricing of average trade size with future returns as shown by the high-low coefficient of 0.207% (t-statistic of 1.07). However, the 100 to 499 and 500 to 999 share grouping does show significance, but in an opposite direction. The 100 to 499 share category shows evidence of positive pricing ability earning 0.487% per month alpha, but with a t-statistic of only 2.27. The 500-999 share trade category earns an alpha of -0.621% per month, or -7.45% per annum, with a robust t-statistic of 3.2. None of the remaining trade size categories demonstrate any significance with future returns.

It should be noted that the robustness of the pricing for the medium trades comes from the low trade size decile, while for the small trades it comes from the high trade size decile. This indicates that in the pre-decimalization period investors would earn positive alphas by investing in stocks that experience a dearth of medium trades, while they would earn positive alphas by investing in high levels of small trades.

The alpha's range from -0.130% to 0.068% across the 1,000 to 9,999 trade sizes to greater than 10,000 share categories with insignificant t-statistics. All of these trade size groupings lack monotonicity across the deciles.

Turning to the post decimalization period, shown in Panel B of Table 5, shows that only the all trade and 500 to 999 share size grouping shows consistency in pricing with future returns. The alphas are -0.499% (t-statistic of 2.68) and -0.369% (t-statistic of 2.30) per month, respectively.

However, the robustness of the t-statistics is reduced relative to the pre-decimalization period with the falloff in pricing of the 500 to 999 share trades potentially a result of the falloff in the number of 500 to 999 share trades subsequent to the decimalization in stock quotes.[8] The significance in pricing for the all trades category seems to rest on the pricing ability of the 500 to 999 share trades.

None of the remaining share trade categories demonstrate any significance, except for the greater than 10,000 share category that shows an alpha of -0.391%, However the significance level is marginal.

The results indicate that only the 500 to 999 share trade category demonstrates consistency in pricing across the pre and post-decimalization periods, although the robustness of the alphas are reduced in the post decimalization period.

As found previously, the robustness in pricing for the medium sized trades mainly derives from the low trade size decile, but a marginally negative alpha is apparent in the top decile. Positive alphas are earned by investing in stocks in the lowest decile of medium trades, and negative alphas are earned investing in high levels of small trades. These results are consistent with retail trading in medium sized trades.

## 8 Trade Size/Trade Direction

I now examine the interaction between trade size and order imbalance. I focus only in the medium sized 500 to 999 share trade size given the prior results that show consistent pricing across the pre and post-decimalization period. I report results in both the pre and post-decimalization period and use Fama-French five-factor (FF-5) alphas, with Newey–West (1987) t-statistics corrections. The intraday trade direction uses the Lee and Ready (1991) methodology from 1983 to 2001 to determine trade direction. Trades above/below the quote midpoint are classified as buys/sells. Trades at the quote midpoint are resolved using the tick test, where buys/sells are classified using the price movement(s) prior to the trade. The tick test is used exclusively from 1980 to 1982 where

[8] Green et al. (2017) notes a significant falloff in cross-sectional pricing ability of anomalies subsequent to 2003. Although not reported,, I find continued pricing of 500 to 999 share trades across to period starting with 2003 and from 2010 indicating a persistence in the pricing results.

only trades are available. The (buys-sells)/(total trades) statistic determines the proportion of buys and sells across quintiles, where quintile one (5) represents extreme sells (buys). I test for pricing using the average trade size against the one-month ahead returns. All results are based on value-weighted.

As shown in Panel A of Table 6, consistent with the results of Panel A of Table 5 the majority of the pricing effect is evident in the low trade size quintiles. Across the lowest trade size quintile, the buy side alpha is 1.0787% per month, while the sell side alpha is 0.44204% per month. The buy-sell alpha is 0.63767% per month. The significance in buy-sell side alphas persist across trade size quintiles one through three, with the bulk of pricing power evident in quintile three. This shows a buy-sell side alpha of 0.83627% per month with the majority of alpha resulting from the buy-side.

The average buy-sell side alpha is 0.44390% per month with a 0.75478% alpha evident for the buys and 0.31088% alpha given for the sells. Both are robustly significant.

It is evident that more informed investors trade in stocks that experience low levels of retail trading. This is consistent with stealth trading (Barclay and Warner, 1993) where more informed investors camouflage their trades by breaking them break down into smaller sizes. Presumably this occurs in small levels of retail trading.

Turning to the post-decimalization period, shown in Panel B of Table 6, illustrates a completely different pricing pattern. Now, negative alphas are most in evidence for small trade size quintiles and for the sell side of the trade. For this period, the sells side show consistently higher positive alphas than the buy side which shows negative alphas.

This is consistent with uninformed trade for the post-decimalization period. It may be that the reduction in number of post-decimalization medium trades disrupted the trading behavior of more informed investors. But this also means that trade size and order flow both evidence retail trading aspects, namely buys (sells) result in negative (positive) alphas and that these effects are concentrated in stocks with low medium trades.

# 9 Controlling for Various Cross-Sectional Effects

Table 7 provides additional analysis to assess the robustness of the pricing results using the 1/0/1 formation/performance strategy with respect to different cross-sectional risk factors. Based on the previously demonstrated outcome, I specifically focus on analyzing the 500 to 999 share portfolio due to its consistent negative alpha in relation to trade size and future returns. The analysis is conducted over the sample period from January 1980 to December 2021 aligns with the availability of firm characteristics.

Table 7 reports FF-5 factor alphas that are independently sorted among the risk variables, with the difference in FF-5 alphas between the quintile portfolios with the highest and lowest 500 to 999 trade size, together with t-statistics to test their statistical significance. The portfolios are value-weighted using lagged firm size as the weight and characteristics as found in Green et al. (2017) and Lewellen (2015). I only report a subset of the results. A full set of results reside in an on-line appendix.

## 9.1 Controlling for Size

I initially control for size by forming quintile portfolios ranked on market capitalization. Then, within each size quintile, I sort stocks into quintile portfolios ranked on trade size. Hence, within each firm size quintile, quintile one contains the stocks with the lowest trade size. The first panel of Table 7 shows that across each firm size quintile, the lowest firm size quintile has a dramatically higher FF-5 alpha than the other quintiles. The lowest trade size quintile alphas range from 0.122% in firm size quintile one to 0.478% in firm size quintile five. The ranking produce monotonic alphas across the firm size quintiles. The 5-1 differences in FF-5 alphas are most pronounced for firm size quintiles three through five ranging from -0.543%, -0.307%, and -0.536% per month, respectively. The t-statistics of these alphas are all above 2.4 in absolute magnitude. In contrast, the 5-1 alphas for the smallest firm size quintiles are statistically insignificant at the 5% level. It is not small stocks that are driving these results. The row labeled " Controlling for Size" averages across the five size quintiles to produce quintile portfolios with dispersion in trade size, but which contain all sizes of firms. The alphas range from 0.275% to -0.034% per month, and after controlling for size, the 5-1 difference in FF-5 alphas is -0.310% per month. Thus, market capitalization does not explain the high returns to low trade size stocks. This is the same procedure followed by Ang, Hodrick, Xing, and Zhang (2006).

In the remainder of Table 7, I repeat the explicit double-sort characteristic controls, replacing size with other stock characteristics. I first form portfolios based on a particular characteristic, then I sort on trade size, and finally I average across the characteristic portfolios. This ensures a dispersion in trade size but contains all aspects of the characteristic.

### 9.2 Controlling for Turnover

Gervais, Kaniel, and Mingelgrin (2001) find that stocks with higher volume have higher returns. Perhaps stocks with low returns are merely stocks with low trading volume given as turnover and measured by trading volume divided by the total number of shares outstanding over the previous month. Turnover is a noisy proxy for liquidity as shown by Datar, Y. Naik, and Radcliffe (1998). It may be the case that stocks with low levels of 500-999 trade size are also stocks with low turnover, hence liquidity may be explaining the results. Table 7 shows that the high alphas on low 500-999 trade size stocks are robust to controlling for turnover. The low trade size quintile earns a significant 0.371% alpha while the high trade size quintile shows a 0.004% alpha. The 5-1 difference in FF-5 alphas is -0.367% per month, and it is highly significant with a t-statistic of -4.54. Although not reported, an examination of the individual turnover quintiles indicates that the 5-1 differences in alphas are most pronounced in the quintile portfolios with the highest, not the lowest, turnover. Similar results are obtained for standardized turnover, defined as the volatility of share turnover (Chordia, Subrahmanyam, and Anshuman, 2001).

### 9.3 Controlling for One-Month Reversal

An alternative liquidity control is the one-month cumulative return that Nagel (2012) argues is a liquidity provision, and is measured as the one-month reversal effect (Jegadeesh and Titman, 1993). In order for the return reversal to be an explanation, low trade size stocks must have a low corresponding returns. Controlling for the one-month cumulative return measure does little to remove the effect. The FF-5 alpha of the lowest trade size portfolio is 0.352%, while the 5-1 difference in alphas is -0.336% with a t-statistic of 4.07. Similar results are found using the Jegadeesh and Titman (1993) six-month momentum where the 5-1 alpha is -0.399% per month with a robust t-statistic of 5.14. This indicates that low trade size quintiles experience relatively high levels of momentum producing a high alpha.

### 9.4 Controlling for Return Volatility

It may be that idiosyncratic volatility (Ang et al., 2006) explains the pricing of trade size. If low trade stocks are associated with high idiosyncratic volatility, then the return volatility effect may explain the pricing of trade size. However, controlling for the idiosyncratic volatility does little to remove the effect. The FF-5 alpha of the lowest trade size portfolio is 0.266%, while the 5-1 difference in alphas is -0.337% with a t-statistic of 3.79.

### 9.5 Controlling for Reversal Effects

Jegadeesh (1990) argues that high prior months returns lead to low future returns. If the reversal effect is to explain the trade size effect, high trade size stocks must have high prior monthly returns. Controlling for the one-month return does not explain the trade size effect with the 5-1 quintile FF-3 alpha recorded at -0.4269 and a t-statistic of 4.64.

### 9.6 Controlling for Momentum Effects

Jegadeesh and Titman (1993) argues that high/low prior month's returns lead to high/low future returns. If the momentum effect is to explain the trade size effect, high trade size stocks must have low prior monthly return. Controlling for the one-month return does not explain the trade size effect with the 5-1 quintile FF-3 alpha recorded at -0.4701 and a t-statistic of 4.90. Similar results are found for the change in momentum.

### 9.7 Controlling for Book-to-Market Ratios

It is generally believed that high book-to-market firms (value firms) have high average returns. Thus, in order for the book-to-market effect to be an explanation of the trade size, the low trade size portfolios must be primarily composed of value stocks that have higher average returns than growth stocks. The row labeled " Controlling for Book-to-Market" shows that across various book-to-market ratios, stocks with the lowest trade size have a robustly significant 0.353% FF-5 alpha. The 5-1 difference in FF-5 alphas is -0.302% per month, with a t-statistic of 3.62.

## 10 Fama-MacBeth Regressions

I further assess the cross-sectional association between one month ahead returns across all share trade categories using a Fama-MacBeth regression controlling for both firm-specific and market characteristics that Green et al. (2017) and Lewellen (2011) show are related to future monthly returns. I run value-weighted regression using the prior month's firm size as the weight. I correct the standard errors using the Newey-West procedure with three lags. I run each Fama-MacBeth regression across the whole period from 1980 to 2021, and if there is significance in the results, I breakdown the period to include 1984 to 2022, from 1989 to 2021, and from 2001 to 2021. This allows for the inclusion of firm characteristics, such as R&D to market capitalization and analyst coverage, that are not available over the whole sample period.

The results are shown in Table 8. As shown, little statistical significance is observed from the 100 to 499 share category that reports a coefficient of -0.0031% with a t-statistic of 1.39. This extends to the 1,000 to 4,999 trade size category that shows a -0.0252% coefficient (1.62 t-statistic). The larger share trade categories demonstrate marginal significance in association with future returns with coefficients of -0.0495% and -0.07866% for the 5,000 to 9,999 share and $\geq$ 10,000 share category, respectively.

However, a relatively robust significance of 2.54 is observed for the 500 to 999 share coefficient across the 1980 to 2021 time period. A one percent increase in lagged 500 to 999 share trades results in a 0.0265% decline in future returns. The significance in association highlights the independence of the 500 to 999 trade size 'anomaly' with other priced anomalies known to exist. It is also telling that of the 22 most prominent anomalies, only eight display any statistical significance with future returns.

Moving to the period 1984 to 2021 that allows for analyst coverage shows slightly diminished significance in the association between 500 to 999 share trades and future returns with the t-statistic falling to 2.37. Again these results show independence with the 22 other priced anomalies, where only seven are now significantly related to future returns. Including R&D expenses that are available from 1989 to 2022 shows continued significance of trade size with future returns as given by a t-statistic of 2.59.

Finally running the Fmama-MacBeth procedure from 2001 to 2021 shows remarkable significance in trade size with a t-statistic of 2.26. The rise in magnitude of the coefficient is consistent with the fall-off in the number of 500 to 999 share trades across the period 2001 to 2021. But, the significance levels are consistent with prior periods illustrating that even in the post decimalization period where it is widely know that the pricing ability of known anomalies is dissipated, the pricing ability of trade size remains significant. Note also that only five of the 22 anomalies are remotely significant subsequent to decimalization.

## 11 Earnings Announcements and Trade Size

I now test the association of trade size using the earnings announcements as an information event. This will provide a test of association with trade size against a known information event. I present Fama and MacBeth (1973) regressions based on the Cumulative Abnormal Return (CAR) surrounding the Earnings Announcement Dates (EAD). The analysis uses CARs from day -1 to +1 relative to the EAD and as the dependent variable and trade size as the independent variable. All regressions are equally-weighted across al NYSE/Amex/NYSE Arca firms between 1980 and 2021. To form distinct portfolios, trades are categorized based on established trade size clusters: 100 to 499 shares, 500 to 999 shares, 1000 to 4999 shares, 5000 to 9999 shares, and finally, trades of $\geq$ 10000 shares. I exclude the three days preceding the EAD and calculate the 15-day moving average (-3 to -18) to determine the percentage of trades within each trade size category. CAR's are determined using a five-factor Fama-French model. I further divide the period into two sub-periods: 1980 to 2000 as the pre-decimalization period and 2001 to 2021 as the post-decimalization period. I further subset the results if significance is observed across the 1980 to 2021 period. The findings are reported in Table 9.

As shown in Table 9, only the medium sized trades are associated with the information content of earnings. Small trades demonstrate a positive, but insignificant, 0.0052 coefficient, while the larger trades are essentially zero. The opposite is found for the medium trades.

For these retail trades, there is a negative and significant relation between CAR and trade size. Higher levels of CAR are associated with lower retail trading, as embodied in 500 to 999 or 1,000 to 4,999 share trades. The coefficients are -0.0017 and -0.0011, respectively. This shows that higher (lower) levels in CAR's draw lower (higher) levels of retail trading. This would be consistent with these retail traders being uninformed about the information content in earnings. This result reinforces the prior pricing and cross-section Fama and MacBeth (1973) tests.

Breaking down the tests into pre and post-decimalization shows that only the 500 to 999 medium trades are reliably associated with CARs across either period. The t-statistics for the 500 to 999 share trades are all above the 5% level, while those between 1,000 and 4,999 shares are weakly significant at only the 10% level. Again higher (lower) levels of 500 to 999 share (medium) trades are associated with lower (higher) CAR levels.

In essence, more informed investors should seek out stocks with high (low) levels of standardized unexpected returns (SUE) that predict the firm's CARs, while avoiding those with high (low) levels of medium trades.

## 12 Conclusions

In this paper I provide evidence that medium sized trades, embodied by trades between 500 and 999 shares, display characteristics consistent with retail trading. Linking prior retail trading proxies such as off-exchange retail orders and relatively low dollar trades, I show a clear association between medium sized trades and these retail trading proxies.

Importantly, stocks with higher (lower) levels of medium trades earn significantly lower (higher) future returns than otherwise similar stocks. The cross-sectional alpha is a negative 7% per annum and is persistent across the pre and post-decimalization periods that are well known to affect the pricing ability of many anomalies. In essence, it is profitable to trade against the prevalence of medium trades. The cross-sectional pricing of trade size is independent to a host of known anomalies in cross-sectional sort and Fama and MacBeth (1973) regression tests.

An important advantage of using medium sized trades is that it is based on widely available intraday transaction data. All that is required is access to TAQ without the need to link the trades with the bid and ask quotes. This approach can more easily determine if behavioral biases are present across a long time. The identification of medium sized trades can also be instrumental in discerning whether there are seasonal and time-series variation in pricing of retail trades again across long time periods. As noted by Hvidkjaer (2008), small trades are most probably a piece of a large institutional order after the decimalization period in 2001 and Reg NMS in 2005. As a result, many, if not all, papers avoid using an analysis of retail/institutional trading in U.S. equity markets.

## References


Alexander, G., and M. Peterson. 2007. An analysis of trade-size clustering and its relation to stealth trading. *Journal of Financial Economics.*

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The Cross-Section of Volatility and Expected Returns. *Journal of Finance.* 61(1):259–299.

Barber, B. M., T. Odean, and N. Zhu. Jan. 2009. Do Retail Trades Move Markets?. *The Review of Financial Studies.* 22(1):151–186.

Barclay, M., and J. Warner. 1993. Stealth Trading and Volatility. *Journal of Financial Economics.* 34(2):281–305.

Battalio, R. H., and R. R. Mendenhall. 2005. Earnings expectations, investor trade size, and anomalous returns around earnings announcements. *Journal of Financial Economics.* 77:289–319.

Boehmer, E., C. Jones, X. Zhang, and X. Zhang. Oct. 2021. Tracking Retail Investor Activity. *Journal of Finance.* LXXVI(5):2249–2306.

Brandt, M. W., A. Brav, J. R. Graham, and A. Kumar. 2010. The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes?. *Review of Financial Studies.* 23(2):863–899.

Brennan, M. J., and A. Subrahmanyam. jan 1998. The Determinants of Average Trade Size. *The Journal of Business.* 71(1):1–25.

Campbell, J. Y., T. Ramadorai, and A. Schwartz. apr 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics.* 92(1):66–91.

Chakravarty, S.. 2001. Stealth-trading: Which traders' trades move stock prices?. *Journal of Financial Economics.* 61(2):289–307.

Chordia, T., and A. Subrahmanyam. 2004. Order imbalance and individual stock returns: Theory and evidence. *Journal of Financial Economics.* 72(3):485–518.

Chordia, T., A. Subrahmanyam, and V. R. Anshuman. 2001. Trading activity and expected stock returns. *Journal of Financial Economics*.

Cready, W. M.. 1988. Information Value and Investor Wealth: The Case of Earnings Announcements. *Journal of Accounting Research*. 26(1):1.

Cready, W. M., and P. G. Mynatt. Apr. 1991. The Information Content of Annual Reports: A Price and Trading Response Analysis. *Accounting Review*. 66(2):291–312.

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

Datar, V. T., N. Y. Naik, and R. Radcliffe. 1998. Liquidity and stock returns: An alternative test. *Journal of Financial Markets*.

Ellis K., M. R., and O. M.. Dec. 2000. The Accuracy of Trade Classification Rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis*. 35(4):529–551.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of political economy*. 81(3):607–636.

Gervais, S., R. Kaniel, and D. H. Mingelgrin. 2001. The High Volume Return Premium. *Journal of Finance*. 56:877–919.

Gompers, P., and R. Metrick. 2001. Institutional Investors and Stock Prices. *Quarterly Journal of Economics*. 166:229–259.

Green, J., J. R. Hand, and X. F. Zhang. Dec. 2017. The characteristics that provide independent information about average u.s. monthly stock returns. *Review of Financial Studies*. Volume 30 (12):4389–4436. https://doi.org/10.1093/rfs/hhx019.

Grinblatt, M., and S. Titman. 1989. Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies*. 2:393–416.

Harvey, C. R., Y. Liu, and H. Zhu. oct 2015. . . . and the Cross-Section of Expected Returns. *Review of Financial Studies*. 29(1):5–68.

Hou, K., C. Xue, and L. Zhang. Sept. 2015. Digesting Anomalies: An Investment Approach. *Review of Financial Studies*. pages 650–705.

Huh, S.-W., and A. Subrahmanyam. 2005. Order Flow Patterns around Seasoned Equity Offerings and their Implications for Stock Price Movements*. *International Review of Finance*. 5(1-2): 75–111.

Hvidkjaer, S.. 2006. A trade-based analysis of momentum. *Review of Financial Studies*. 19(2): 457–491.

——— . 2008. Small Trades and the Cross-Section of Stock Returns. *Review of Financial Studies*. 21(3):1123–1151.

Jegadeesh, N.. 1990. Evidence of Predictable Behavior of Security Returns. *Journal of Finance*. 45 (3):881–898.

Jegadeesh, N., and S. Titman. 1993. Returns to Buying Winners and Selling Losers; Implications for Stock Market Efficiency. *Journal of Finance*. 48:65–91.

Kaniel, R., S. Liu, G. Saar, and S. Titman. 2012. Individual Investor Trading and Return Patterns around Earnings Announcements. *Journal of Finance*. 67:639–680.

Kaniel, R., G. Saar, and S. Titman. 2008. Individual Investor Trading and Stock Returns. *Journal of Finance*. LXIII, No. 1:273–310.

Kelley, E. K., and P. C. Tetlock. 2013. How Wise Are Crowds? Insights from Retail Orders and Stock Returns. *Journal of Finance*. 68:1229–1265.

Lee, C. M.. 1992. Earnings News and Small Traders. *Journal of Accounting and Economics*. 15: 265–302.

Lee, C. M., and B. Radhakrishna. 2000. Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets*. 3(2):83 – 111.

Lee, C. M., and M. J. Ready. 1991. Inferring Trade Direction from Intraday Data. *Journal of Finance*. 46, No. 2:733–746.

Lesmond, D., J. Ogden, and C. Trzcinka. 1999. A New Estimate of Transaction Costs. *The Review of Financial Studies*. 5:1113–1141.

Lewellen, J.. oct 2011. Institutional investors and the limits of arbitrage. *Journal of Financial Economics*. 102(1):62–80.

——— . jun 2015. The Cross-section of Expected Stock Returns. *Critical Finance Review*. 4(1): 1–44.

Malmendier, U., and D. Shanthikumar. Do Security Analysts Speak in Two Tongues? Forthcoming, Review of Financial Studies. 2014.

Nagel, S.. 2012. Evaporating Liquidity. *The Review of Financial Studies*. Vol. 25(No. 7 (July 2012)): 2005–2039.

Newey, W., and K. West. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*. 55:703–708.

Odean, T.. 1998. Volume, volatility, price, and profit when all traders are above average. *Journal of Finance*.

**Table 1:** Summary Statistics

Trade size statistics are presented across the pre-decimalization period, 1979m1 to 2000m12, and post-decimalization period, 2001m1 to 2021m12. Trades are grouped into separate portfolios using established (Barclay and Warner, 1993) trade size clusters between 100 to 499 shares, 500 to 999 shares, 1000 to 4999 shares, 5000 to 9999 shares, and finally $\geq$ 10000 shares. These trade size clusters are formed using all daily trades within each trade size grouping and then divided by the total number of trades per day. This total is then averaged across the month. For each period I show the mean, median, standard deviation, minimum, and maximum values for each trade size.

| | Panel A: Pre-Decimalization Period 1979m1 to 2000m12 | | | | | |
|---|---|---|---|---|---|---|
| | Trades | 100 to 499 | 500 to 999 | 1000 to 4999 | 5000 to 9999 | $\geq$ 10000 |
| Mean | 78.15 | 54.72 | 18.81 | 21.84 | 2.61 | 1.97 |
| Median | 55.08 | 51.06 | 19.29 | 24.30 | 3.04 | 2.14 |
| Std.Dev. | 64.24 | 8.67 | 1.67 | 5.41 | 0.93 | 0.74 |
| Min | 18.32 | 44.49 | 14.34 | 11.14 | 0.84 | 0.50 |
| Max | 327.15 | 72.88 | 21.39 | 27.97 | 3.83 | 3.12 |
| | Panel B: Post-Decimalization Period 2001m1 to 2021m12 | | | | | |
| Mean | 6101.80 | 71.06 | 6.50 | 5.74 | 0.62 | 0.40 |
| Median | 6870.20 | 69.70 | 3.52 | 2.32 | 0.20 | 0.14 |
| Std.Dev. | 3855.56 | 17.74 | 5.93 | 6.39 | 0.70 | 0.45 |
| Min | 323.31 | 33.28 | 1.40 | 0.94 | 0.09 | 0.05 |
| Max | 19081.26 | 94.89 | 20.52 | 26.67 | 3.26 | 2.20 |

**Table 2:** Trade Size Determinant Tests

The table reports Fama-MacBeth cross-sectional determinant regressions across all NYSE/Amex/NYSE Arca firms across the period 1980 to 2022 using trade size clusters between 100 and 499 shares, 500 and 999 shares, 1000 and 4999 shares, 5000 and 9999 shares, and trades $\geq$ 10000 shares. I use the lagged trade size to test for persistence. I also include the past one-month return for the liquidity provision and the 11-month momentum return along with the control variables of firm size, the percentage of institutional holdings, the number of institutional holders (owners), book-to-market, return volatility, and turnover. The control variables are log scaled, except for the momentum variables. T-statistics are presented in the parentheses. Significance is reported with an * (10% significance), an ** (5% significance), or an *** (1% significance).

| Variable | 100 to 499 | 500 to 999 | 1000 to 4999 | 5000 to 9999 | $\geq$ 10000 |
|---|---|---|---|---|---|
| 1-Month Momentum | 0.0289*** | -0.0073*** | -0.0241*** | -0.0032*** | -0.0031*** |
| | (8.89) | (-6.35) | (-13.48) | (-7.93) | (-7.99) |
| 11-Month Momentum | 0.0016*** | -0.0012*** | -0.0025*** | -0.0003*** | -0.0000 |
| | (2.80) | (-4.24) | (-6.98) | (-2.77) | (-0.45) |
| % Institutional Holdings | -0.0029*** | -0.0009*** | 0.0003 | 0.0012*** | 0.0011*** |
| | (-6.15) | (-3.72) | (0.99) | (8.91) | (9.45) |
| Number of Owners | 0.0003 | -0.0014*** | 0.0012** | -0.0001 | -0.0002 |
| | (0.38) | (-2.62) | (2.15) | (-0.42) | (-1.20) |
| Book-to-Market | 0.0006** | -0.0003** | -0.0002 | 0.0001*** | 0.0003*** |
| | (2.24) | (-2.31) | (-1.06) | (2.63) | (6.73) |
| Turnover | -0.0012** | 0.0019*** | 0.0027*** | 0.0007*** | 0.0000 |
| | (-2.18) | (7.14) | (8.52) | (8.35) | (0.66) |
| Return Volatility | -0.0093*** | 0.0034*** | 0.0076*** | 0.0015*** | 0.0012*** |
| | (-11.71) | (7.59) | (11.81) | (10.25) | (9.69) |
| Firm Size | -0.0015*** | -0.0010*** | 0.0007* | 0.0010*** | 0.0009*** |
| | (-2.74) | (-2.67) | (1.94) | (9.08) | (8.35) |
| Lag(100 to 499 Shares) | 0.8115*** | | | | |
| | (85.36) | | | | |
| Lag(500 to 999 Shares) | | 0.6063*** | | | |
| | | (34.55) | | | |
| Lag(1000 to 4999 Shares) | | | 0.7392*** | | |
| | | | (76.00) | | |
| Lag(5000 to 9999 Shares) | | | | 0.5890*** | |
| | | | | (37.02) | |
| $\geq$ 10000 Shares | | | | | 0.6574*** |
| | | | | | (47.64) |
| cons | 0.0835*** | 0.0974*** | 0.0607*** | 0.0032*** | -0.0003 |
| | (12.85) | (12.49) | (10.96) | (4.47) | (-0.51) |
| | | | | | (50.08) |
| Intercept | 0.0817*** | 0.1010*** | 0.0556*** | 0.0024*** | -0.0007 |
| | (13.43) | (11.80) | (10.55) | (3.46) | (-1.57) |
| Adjusted $R^2$ | 74.49 | 49.70 | 65.27 | 48.35 | 54.13 |

**Table 3:** Retail Trading Decomposition by Trade Size

Panel A defines retail trades are defined as the percentage of trades less than $10,000 (Lee, 1992; Brandt et al., 2010), for the period 1980 to 2000. Panel B defines retail trades based on TRF trades shown in Boehmer et al. (2021) for the period 2010 to 2015. There are 252 monthly observations for Panel A. Panel B is further separated into weekly returns (top panel, 312 observations) and then monthly returns (bottom panel, 72 observations). Trades are grouped into separate portfolios using established (Barclay and Warner, 1993) trade size clusters between 100 to 499 shares, 500 to 999 shares, 1000 to 4999 shares, 5000 to 9999 shares, and finally $\geq$ 10000 shares. Panel A uses a cross-sectional Fama-MacBeth regression using the the percentage of retail trades as the dependent variable and various trade size categories as the independent variable. Panel B uses quintile sorts on one-week or one-month ahead returns with order-imbalance and with alphas determined via a FF-3 factor model. Newey and West (1987) t-statistics corrected with 5 lags are presented in parentheses. Significance is reported with an * (10% significance), an ** (5% significance), or an *** (1% significance).

| | Panel A: Small Trades Retail Trading | | | | |
|---|---|---|---|---|---|
| 100 to 499 Shares | -0.1371** | | | | |
| | (-2.15) | | | | |
| 500 to 999 Shares | | 0.3137*** | | | |
| | | (4.95) | | | |
| 1000 to 4999 Shares | | | 0.2044*** | | |
| | | | (2.60) | | |
| 5000 to 9999 Shares | | | | -1.0560** | |
| | | | | (-2.48) | |
| $\geq$ 10000 Shares | | | | | -1.7554*** |
| | | | | | (-5.73) |
| Intercept | 0.2154*** | 0.1085*** | 0.0940*** | 0.1895*** | 0.2012*** |
| | (6.36) | (12.36) | (4.75) | (15.47) | (23.65) |
| Adjusted $R^2$ | 4.38 | 1.55 | 4.82 | 4.86 | 5.92 |

| | Panel B: Off-Exchange Trade Reporting Facility Retail Trades | | | | | |
|---|---|---|---|---|---|---|
| Trade Category | Sells (1) | 2 | 3 | 4 | Buys (5) | Buys-Sells (5 - 1) |
| | | | Weekly Returns | | | |
| All Trades | 0.267 | 0.294 | 0.292 | 0.344 | 0.373 | 0.106*** |
| | (2.86) | (3) | (2.98) | (3.54) | (4.04) | (2.72) |
| 100 to 499 Shares | 0.258 | 0.293 | 0.297 | 0.330 | 0.391 | 0.133*** |
| | (2.62) | (3.03) | (3) | (3.42) | (4.25) | (3.25) |
| 500 to 999 Shares | 0.307 | 0.272 | 0.311 | 0.325 | 0.406 | 0.099*** |
| | (3.17) | (2.9) | (3.15) | (3.35) | (4.12) | (3.38) |
| 1000 to 4999 Shares | 0.349 | 0.264 | 0.284 | 0.348 | 0.416 | 0.065** |
| | (3.51) | (2.88) | (2.94) | (3.51) | (3.99) | (2.37) |
| | | | Monthly Returns | | | |
| All Trades | 0.183 | -0.081 | 0.057 | 0.047 | 0.077 | -0.106 |
| | (1.57) | (-1.07) | (1.1) | (0.65) | (0.37) | (-0.43) |
| 100 to 499 Shares | 0.309 | -0.085 | 0.075 | 0.006 | 0.009 | -0.300* |
| | (3.02) | (-1.05) | (1.44) | (0.09) | (0.05) | (-1.74) |
| 500 to 999 Shares | 0.365 | -0.002 | 0.056 | -0.087 | 0.131 | -0.234* |
| | (3.42) | (-0.03) | (1.05) | (-0.95) | (1.54) | (-1.75) |
| 1000 to 4999 Shares | 0.325 | -0.045 | 0.019 | -0.026 | 0.150 | -0.175 |
| | (3.27) | (-0.5) | (0.32) | (-0.27) | (1.87) | (-1.34) |

**Table 4:** Decile Sorts of Trade Size Groups

Value-weighted five-factor Fama-French alphas are presented using one-month ahead returns, sorted into deciles based on trade size portfolios for NYSE/Amex/NYSE Arca firms from 1980 to 2022. The prior month's market capitalization is used as the weight. The established (Barclay and Warner, 1993) trade size clusters of 100-499, 500-999, 1000-4999, 5000-9999, and $\geq$ 10000 shares are used to group trades into separate portfolios. These trade size clusters are formed using all daily trades within each trade size grouping and then divided by the total number of trades per day. This total is then averaged across the month. The 'Low' ('High') portfolio consists of stocks with the lowest (highest) concentration of share trades for each trade size cluster. Only firms with a prior month share price >= $5.00 are included, and trades above or below a round penny are deleted from 2006 to 2017 (Boehmer et al., 2021). The column High-Low shows the long-short portfolio across deciles. The Newey and West (1987) t-statistics, corrected with 3 lags, are presented in parentheses, with significance denoted by * (10% significance), ** (5% significance), or *** (1% significance).

| Trade Group | 1 (Low) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 (High) | High - Low |
|---|---|---|---|---|---|---|---|---|---|---|---|
| All Trades | 0.389 | 0.422 | 0.207 | 0.274 | 0.191 | 0.286 | 0.200 | 0.202 | 0.218 | 0.153 | -0.237 |
| | (2.66) | (3.48) | (2.1) | (3.41) | (2.66) | (3.74) | (2.66) | (2.82) | (3.63) | (3.81) | (-1.59) |
| 100 to 499 Shares | 0.079 | 0.197 | 0.257 | 0.116 | 0.281 | 0.344 | 0.399 | 0.369 | 0.325 | 0.330 | 0.251 |
| | (0.8) | (2.62) | (3.89) | (1.68) | (4.29) | (4.75) | (5.48) | (4.14) | (3.22) | (3.1) | (1.55) |
| 500 to 999 Shares | 0.459 | 0.402 | 0.376 | 0.270 | 0.229 | 0.169 | 0.098 | 0.163 | 0.006 | -0.171 | -0.631*** |
| | (5.08) | (4.99) | (4.8) | (3.75) | (3.14) | (2.45) | (1.23) | (1.99) | (0.07) | (-1.76) | (-4.8) |
| 1000 to 4999 Shares | 0.354 | 0.357 | 0.422 | 0.322 | 0.276 | 0.341 | 0.240 | 0.176 | 0.133 | -0.037 | -0.391** |
| | (3.26) | (3.98) | (4.68) | (4.56) | (3.57) | (4.66) | (3.6) | (2.47) | (1.74) | (-0.35) | (-2.44) |
| 5000 to 9999 Shares | 0.475 | 0.185 | 0.091 | 0.223 | 0.319 | 0.309 | 0.339 | 0.205 | 0.060 | 0.060 | -0.415*** |
| | (3.46) | (2) | (0.79) | (2.86) | (4.04) | (4.55) | (4.83) | (3.11) | (0.78) | (0.73) | (-2.63) |
| $\geq$ 10000 Shares | 0.373 | 0.283 | 0.022 | 0.282 | 0.173 | 0.291 | 0.233 | 0.216 | 0.100 | 0.131 | -0.242 |
| | (2.89) | (1.87) | (0.25) | (3.61) | (2.12) | (4.46) | (3.83) | (3.46) | (1.32) | (1.46) | (-1.54) |

**Table 5:** Pre/Post Decimalization Sort Tests

Value-weighted five-factor Fama-French alphas using one-month ahead returns are presented using decile sorts across the pre-decimalization period from 1980m1 to 2000m12 and the post-decimalization period from 2001m1 to 2022m12 for NYSE/Amex/NYSE Arca firms. Trades are grouped into separate portfolios using established (Barclay and Warner, 1993) trade size clusters between 100 to 499 shares, 500 to 999 shares, 1000 to 4999 shares, 5000 to 9999 shares, and finally $\geq$ 10000 shares. These trade size clusters are formed using all daily trades within each trade size grouping and then divided by the total number of trades per day. This total is then averaged across the month. The "Low" ("High") portfolio contains those stocks with the lowest (highest) concentration of share trades for each trade size cluster. The prior month's firm size is used as the weight. Only firms with a prior month share price >= $5.00 are included in the analysis. I also delete those trades above or below a round penny (Boehmer et al., 2021) from 2007 to 2017. The column High-Low refers to the long-short portfolio across the deciles. Newey and West (1987) t-statistics corrected with 3 lags are presented in parentheses. Significance is reported with an * (10% significance), an ** (5% significance), or an *** (1% significance).

| Trade Group | 1 (Low) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 (High) | High - Low |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pre-Decimalization - 1980m1 to 2000m12 | | | | | | | | | | | |
| All Trades | 0.198 | 0.500 | 0.264 | 0.253 | 0.147 | 0.265 | 0.236 | 0.207 | 0.284 | 0.406 | 0.207 |
| | (1.02) | (3.09) | (1.65) | (1.99) | (1.32) | (2.22) | (2.09) | (1.9) | (3.13) | (8.14) | (1.07) |
| 100- 499 Shares | 0.192 | 0.260 | 0.443 | 0.164 | 0.440 | 0.614 | 0.467 | 0.615 | 0.581 | 0.679 | 0.487** |
| | (1.47) | (2.12) | (4.43) | (1.81) | (5.04) | (6.68) | (5.02) | (4.48) | (4.24) | (4.43) | (2.27) |
| 500- 999 Shares | 0.585 | 0.628 | 0.524 | 0.410 | 0.225 | 0.315 | 0.186 | 0.306 | 0.044 | -0.025 | -0.610*** |
| | (4.83) | (5.95) | (5.81) | (4.06) | (2.03) | (3.2) | (1.62) | (2.38) | (0.31) | (-0.17) | (-3.2) |
| 1000- 4999 Shares | 0.430 | 0.518 | 0.650 | 0.476 | 0.441 | 0.490 | 0.352 | 0.331 | 0.357 | 0.300 | -0.130 |
| | (2.59) | (3.52) | (4.56) | (4.87) | (4.9) | (4.83) | (3.99) | (3.41) | (3.17) | (1.87) | (-0.5) |
| 5000- 9999 Shares | 0.660 | 0.216 | 0.048 | 0.312 | 0.626 | 0.503 | 0.406 | 0.323 | 0.299 | 0.351 | -0.309 |
| | (3.49) | (1.46) | (0.21) | (2.52) | (5.17) | (4.76) | (4.72) | (4.03) | (2.88) | (2.99) | (-1.43) |
| $\geq$ 10000 Shares | 0.385 | 0.624 | -0.113 | 0.510 | 0.311 | 0.515 | 0.277 | 0.352 | 0.264 | 0.453 | 0.068 |
| | (2.34) | (1.92) | (-0.67) | (3.91) | (2.54) | (4.98) | (3.04) | (4.36) | (2.69) | (3.39) | (0.33) |
| Post-Decimalization - 2001m1 to 2022m12 | | | | | | | | | | | |
| All Trades | 0.541 | 0.355 | 0.371 | 0.307 | 0.125 | 0.29 | 0.208 | 0.21 | 0.212 | 0.053 | -0.499*** |
| | (3.24) | (2.25) | (2.7) | (3.45) | (1.18) | (3.55) | (2.55) | (2.66) | (2.27) | (1.56) | (-2.68) |
| 100- 499 Shares | 0.166 | -0.005 | 0.192 | 0.193 | 0.24 | 0.113 | 0.137 | 0.035 | 0.103 | -0.101 | -0.267 |
| | (1.21) | (-0.04) | (2.31) | (2.42) | (2.59) | (1.23) | (1.71) | (0.36) | (0.85) | (-0.49) | (-0.87) |
| 500- 999 Shares | 0.238 | 0.065 | 0.13 | -0.01 | 0.266 | 0.139 | 0.215 | 0.283 | 0.054 | -0.131 | -0.369** |
| | (2.02) | (0.76) | (2.17) | (-0.09) | (2.89) | (1.52) | (2.04) | (1.99) | (0.69) | (-1.42) | (-2.3) |
| 1000- 4999 Shares | -0.017 | 0.074 | 0.041 | 0.084 | 0.19 | 0.27 | 0.194 | 0.346 | -0.01 | -0.11 | -0.093 |
| | (-0.1) | (0.58) | (0.52) | (0.82) | (2.25) | (2.81) | (2.16) | (3.53) | (-0.13) | (-1.15) | (-0.4) |
| 5000- 9999 Shares | 0.093 | -0.061 | 0.042 | 0.079 | 0.217 | 0.332 | 0.06 | 0.306 | -0.2 | -0.173 | -0.248 |
| | (0.63) | (-0.46) | (0.39) | (0.96) | (2.19) | (4.11) | (0.46) | (2.77) | (-1.58) | (-1.53) | (-1.27) |
| $\geq$ 10000 Shares | 0.146 | 0.207 | 0.091 | 0.11 | 0.088 | 0.048 | -0.008 | 0.057 | 0.099 | -0.235 | -0.391* |
| | (0.84) | (1.61) | (0.62) | (1.21) | (0.96) | (0.53) | (-0.08) | (0.57) | (0.83) | (-1.83) | (-1.86) |

## **Table 6:** Trade Size/Trade Direction Double Sorts

The table reports Fama-French five-factor (FF-5) alphas, with robust (three lags) Newey–West (1987) t-statistics in parentheses. All the strategies use the 500-999 share trades computed relative to FF-5, but controls for either the buys-sells computed from the intraday trade direction. The intraday trade direction uses the Lee and Ready (1991) methodology from 1983 to 2001 to determine trade direction. Trades above/below the quote midpoint are classified as buys/sells. Trades at the quote midpoint are resolved using the tick test, where buys/sells are classified using the price movement(s) prior to the trade. The tick test is used exclusively from 1980 to 1982 where only trades are available. The (buys-sells)/(total trades) statistic determines the proportion of buys and sells across quintiles, where quintile one (5) represents extreme sells (buys). The column "High-Low" refers to the difference in FF-5 alphas between portfolio 5 and portfolio 1 for either the rank on 500 to 999 shares or for trade direction. I count all daily 500 to 999 share trades and divide this count by the total number of trades per day. I then average this daily ratio across the month. Trade direction counts the daily number of (buys - sells)/total trades and then averages them across the month. I test for pricing using the average trade size against the one-month ahead returns. All results are based on value-weighted using the lagged month's firm size as the weight deleting any firm with a prior month share price < \$5.00. Significance reported with an * (10% significance), an ** (5% significance), or an *** (1% significance).

| | Panel A: Pre-Decimalization | | | | | |
|---|---|---|---|---|---|---|
| | Low (Sells) | 2 | 3 | 4 | High (Buys) | Buys-Sells |
| Trade Size Rank | | | | | | |
| 1 (Low) | 0.44104 | 0.61524 | 0.58107 | 0.98065 | 1.07870 | 0.63767* |
| | (2.87) | (3.94) | (3.52) | (6.90) | (5.65) | (2.28) |
| 2 | 0.35543 | 0.43965 | 0.47196 | 0.68972 | 0.79236 | 0.43693* |
| | (2.55) | (3.33) | (5.55) | (5.81) | (5.51) | (1.94) |
| 3 | 0.01645 | 0.31464 | 0.43063 | 0.52209 | 0.85273 | 0.83627* |
| | (0.10) | (2.11) | (3.40) | (4.47) | (4.38) | (3.33) |
| 4 | 0.46764 | 0.41935 | 0.43008 | 0.50237 | 0.65532 | 0.18768 |
| | (3.31) | (2.77) | (2.89) | (3.74) | (4.18) | (0.91) |
| 5 (High) | 0.27387 | 0.38797 | 0.20194 | 0.27886 | 0.39483 | 0.12096 |
| | (1.72) | (1.93) | (1.16) | (1.67) | (2.46) | (0.60) |
| High - Low | 0.31088* | 0.43537* | 0.42314* | 0.59474* | 0.75479* | 0.44390* |
| | (4.23) | (5.93) | (5.70) | (8.57) | (9.05) | (3.91) |

| | Panel B: Post-Decimalization | | | | | |
|---|---|---|---|---|---|---|
| | Low (Sells) | 2 | 3 | 4 | High (Buys) | Buys-Sells |
| Trade Size Rank | | | | | | |
| 1 (Low) | 0.32780 | 0.32089 | 0.10158 | 0.34599 | -0.29526 | -0.61588 |
| | (1.79) | (2.55) | (0.92) | (3.06) | (-1.26) | (-1.66) |
| 2 | 0.50250 | 0.16882 | 0.06287 | -0.00647 | -0.02119 | -0.52370 |
| | (5.19) | (1.43) | (0.65) | (-0.06) | (-0.13) | (-2.54) |
| 3 | -0.04622 | 0.17462 | 0.33879 | 0.34949 | -0.04857 | -0.00235 |
| | (-0.25) | (1.10) | (2.47) | (2.85) | (-0.28) | (-0.01) |
| 4 | 0.21194 | 0.16705 | 0.38958 | 0.22609 | -0.11074 | -0.33443 |
| | (0.86) | (1.12) | (3.44) | (2.28) | (-0.49) | (-0.85) |
| 5 (High) | 0.56315 | -0.00414 | 0.08625 | -0.07372 | -0.00973 | -0.58577 |
| | (2.68) | (-0.02) | (0.78) | (-0.57) | (-0.05) | (-1.84) |
| High - Low | 0.31202* | 0.16546* | 0.19680* | 0.16828* | -0.09742 | -0.41249 |
| | (3.45) | (2.47) | (3.51) | (3.01) | (-1.07) | (-2.88) |

## **Table 7:** Alphas of Trade Size in Double Sorts Tests

The table reports Fama-French five-factor alphas with robust (3 lags) Newey and West (1987) t-statistics across NYSE/Amex/NYSE Arca firms. The column "High-Low" refers to the difference in FF-5 alphas between portfolio 5 and portfolio 1. In the panel labeled "Size Quintiles," stocks are independently sorted into five quintiles on the basis of size and by 500-999 share concentrations each month. For the remaining characteristics, similar independent double sort tests are reported, but only for aggregate results. Each month, I independently sort stocks based on the first characteristic and based on the 500-999 trade size concentration relative to the FF-5 model. The five 500-999 trade size portfolios are then averaged over each of the five characteristic portfolios. Hence, they represent 500-999 trade size quintile portfolios controlling for the characteristic. Each control variable stems from the characteristics of Green et al. (2017) and Lewellen (2015). The sample period is 1980 to 2022 and all portfolios are value-weighted using lagged firm size as the weight.

| | Ranking on 500-999 Shares | | | | | |
|---|---|---|---|---|---|---|
| Control Variable | 1 (Low) | 2 | 3 | 4 | 5 (High) | High-Low |
| Size Quintiles: Small | 0.122 | 0.024 | 0.193 | 0.062 | 0.150 | 0.027 |
| | (1.02) | (0.20) | (1.41) | (0.51) | (1.15) | (0.19) |
| 2 | 0.162 | 0.236 | 0.124 | -0.0241 | -0.028 | -0.190 |
| | (1.83) | (2.61) | (1.23) | (-0.20) | (-0.16) | (-1.10) |
| 3 | 0.298 | 0.223 | 0.169 | 0.014 | -0.244 | -0.543 |
| | (3.67) | (2.46) | (1.95) | (0.14) | (-1.93) | (-3.81) |
| 4 | 0.313 | 0.343 | 0.052 | 0.054 | 0.005 | -0.307 |
| | (3.32) | (3.83) | (0.64) | (0.57) | (0.05) | (-2.42) |
| Large | 0.478 | 0.345 | 0.20520 | 0.178 | -0.058 | -0.536 |
| | (5.03) | (5.09) | (3.22) | (2.31) | (-0.53) | (-3.65) |
| Controlling for Firm Size | 0.275 | 0.234 | 0.148 | 0.057 | -0.034 | -0.310 |
| | (5.54) | (4.55) | (2.84) | (1.02) | (-0.54) | (-4.59) |
| Controlling for Turnover | 0.371 | 0.322 | 0.142 | 0.074 | 0.004 | -0.367 |
| | (6.36) | (5.98) | (2.45) | (1.17) | (0.06) | (-4.54) |
| Controlling for Turnover Volatility | 0.417 | 0.357 | 0.163 | 0.074 | -0.024 | -0.442 |
| | (7.26) | (7.05) | (2.92) | (1.28) | (-0.36) | (-5.40) |
| Controlling for One-Month Reversal | 0.352 | 0.342 | 0.170 | 0.105 | 0.016 | -0.336 |
| | (6.22) | (6.55) | (3.10) | (1.80) | (0.23) | (-4.07) |
| Controlling for Six-Month Momentum | 0.407 | 0.299 | 0.127 | 0.092 | 0.007 | -0.399 |
| | (6.81) | (5.11) | (2.22) | (1.49) | (0.11) | (-5.14) |
| Controlling for Return Volatility | 0.266 | 0.266 | 0.131 | 0.054 | -0.071 | -0.337 |
| | (3.84) | (4.29) | (1.88) | (0.81) | (-0.96) | (-3.79) |
| Controlling for Book-to-Market | 0.353 | 0.362 | 0.117 | 0.120 | 0.051 | -0.302 |
| | (6.02) | (7.13) | (2.33) | (2.20) | (0.77) | (-3.62) |
| Controlling for Accruals | 0.405 | 0.374 | 0.135 | 0.070 | 0.008 | -0.396 |
| | (7.63) | (7.15) | (2.63) | (1.31) | (0.14) | (-4.78) |

## **Table 8:** Fama-MacBeth Regressions

Value-weighted Fama-MacBeth regressions using lagged firm size as the weight are presented for all NYSE/Amex/NYSE Arca firms using future monthly returns are presented for various trade size categories and firm characteristics from 1980 to 2021. I use the 12 independent characteristics found in Green et al. (2017) in addition to the variables outlined in Lewellen (2015). $R\&D$ to market capitalization is available only from April, 1989, whereas the change in the number of analysts is available from 1984 to 2021 with the remaining data are available from 1980 to 2021. $R\&D$ to market capitalization is set to zero if the value is missing. Book-to-market, firm size, and zero returns are log scaled. The number of observations counts the number of months for the each regression. The adjusted $R^2$ is the average of the cross–sectional adjusted $R^2$'s. Newey and West (1987) robust $t$-statistics with 3 lags are in parentheses and significance at the 1%, 5%, and 10% level is given by ***, ** and *, respectively.

| Variable | 1980-2022 | 1980-2022 | 1984-2021 | 1989-2021 | 2001-2022 | 1980-2022 | | |
|---|---|---|---|---|---|---|---|---|
| Lag(100 to 499 Shares) | -0.0031 | | | | | | | |
| | (-1.39) | | | | | | | |
| Lag(500 to 999 Shares) | | -0.0265** | -0.0271** | -0.0346*** | -0.0435** | | | |
| | | (-2.54) | (-2.37) | (-2.59) | (-2.26) | | | |
| Lag(1000 to 4999 Shares) | | | | | | -0.0252 | | |
| | | | | | | (-1.62) | | |
| Lag(5000 to 9999 Shares) | | | | | | | -0.4950* | |
| | | | | | | | (-1.86) | |
| $\geq$ 10000 Shares | | | | | | | | -0.7866* |
| | | | | | | | | (-1.79) |
| Earnings Announcement Return | 0.0060 | 0.0062 | 0.0058 | 0.0066 | 0.0075 | 0.0065 | 0.0062 | 0.0061 |
| | (1.38) | (1.39) | (1.28) | (1.42) | (1.61) | (1.49) | (1.42) | (1.39) |
| Cash and Cash Equivalents | 0.0083** | 0.0085** | 0.0081** | 0.0082** | 0.0033 | 0.0088** | 0.0088** | 0.0085** |
| | (2.38) | (2.45) | (2.34) | (2.18) | (0.82) | (2.51) | (2.54) | (2.48) |
| Earnings Increases | 0.0005*** | 0.0005*** | 0.0004** | 0.0005*** | 0.0006*** | 0.0005*** | 0.0005*** | 0.0005*** |
| | (3.04) | (2.84) | (2.53) | (2.83) | (2.62) | (2.95) | (2.84) | (2.85) |
| 1-Month Reversal | -0.0253*** | -0.0251*** | -0.0213*** | -0.0192*** | -0.0199** | -0.0256*** | -0.0258*** | -0.0253*** |
| | (-4.29) | (-4.25) | (-3.47) | (-2.87) | (-2.16) | (-4.36) | (-4.38) | (-4.29) |
| Change in 6-Month Momentum | -0.0055*** | -0.0056*** | -0.0061*** | -0.0060*** | -0.0049** | -0.0055*** | -0.0056*** | -0.0056*** |
| | (-3.90) | (-3.96) | (-4.27) | (-4.03) | (-2.51) | (-3.92) | (-3.92) | (-3.98) |
| Share Turnover | -0.0026*** | -0.0023** | -0.0013 | -0.0012 | -0.0003 | -0.0026*** | -0.0026*** | -0.0025*** |
| | (-2.85) | (-2.56) | (-1.44) | (-1.24) | (-0.32) | (-2.81) | (-2.81) | (-2.70) |
| Volatility Share Turnover | 0.0003** | 0.0003* | 0.0000 | 0.0000 | 0.0001 | 0.0003** | 0.0003* | 0.0003* |
| | (2.03) | (1.94) | (0.26) | (0.32) | (0.84) | (2.09) | (1.95) | (1.78) |
| Return Volatility | -0.1700*** | -0.1599** | -0.1647** | -0.1940** | -0.1724 | -0.1616** | -0.1663** | -0.1725*** |
| | (-2.60) | (-2.42) | (-2.36) | (-2.51) | (-1.59) | (-2.48) | (-2.53) | (-2.61) |
| Zero Trades | -0.0002 | -0.0003 | -0.0002 | -0.0002 | 0.0002 | -0.0002 | -0.0002 | -0.0002 |
| | (-0.53) | (-0.70) | (-0.53) | (-0.33) | (0.25) | (-0.56) | (-0.53) | (-0.60) |
| R&D to Market Cap. | | | | 0.0166 | 0.0176 | | | |
| | | | | (1.14) | (1.04) | | | |
| Analyst Change | | | -0.0000 | 0.0000 | -0.0001 | | | |
| | | | (-0.04) | (0.38) | (-0.54) | | | |
| Total Debt to Market Cap. | 0.0000 | 0.0000 | -0.0000 | 0.0001 | -0.0002 | -0.0000 | 0.0000 | -0.0000 |
| | (0.00) | (0.02) | (-0.18) | (0.28) | (-0.91) | (-0.03) | (0.06) | (-0.12) |
| Asset Growth | -0.0010 | -0.0011 | -0.0005 | -0.0007 | 0.0004 | -0.0011 | -0.0012 | -0.0011 |
| | (-0.90) | (-1.01) | (-0.47) | (-0.66) | (0.31) | (-1.00) | (-1.08) | (-1.06) |
| 35-Month Momentum | -0.0004 | -0.0005 | -0.0004 | -0.0010 | -0.0011 | -0.0005 | -0.0005 | -0.0005 |
| | (-0.42) | (-0.60) | (-0.41) | (-0.94) | (-0.79) | (-0.60) | (-0.54) | (-0.52) |
| 11-Month Momentum | 0.0019 | 0.0017 | 0.0008 | 0.0012 | -0.0025 | 0.0017 | 0.0017 | 0.0017 |
| | (0.85) | (0.78) | (0.33) | (0.46) | (-0.73) | (0.75) | (0.76) | (0.77) |
| % Change Shares Outstanding | -0.0029*** | -0.0028*** | -0.0030*** | -0.0030*** | -0.0043*** | -0.0028*** | -0.0028*** | -0.0027*** |
| | (-3.43) | (-3.29) | (-3.30) | (-2.97) | (-2.99) | (-3.37) | (-3.32) | (-3.29) |
| Sales to Market Cap. | 0.0003 | 0.0003 | 0.0003 | 0.0000 | 0.0007* | 0.0003 | 0.0003 | 0.0003 |
| | (1.18) | (1.06) | (0.96) | (0.09) | (1.68) | (1.08) | (1.03) | (1.02) |
| Dividend to Market Cap. | -0.0031 | -0.0028 | 0.0036 | 0.0024 | 0.0333 | -0.0006 | 0.0029 | 0.0004 |
| | (-0.13) | (-0.12) | (0.14) | (0.08) | (1.05) | (-0.03) | (0.12) | (0.02) |
| Accruals | -0.0062 | -0.0066 | -0.0085 | -0.0044 | -0.0044 | -0.0063 | -0.0069 | -0.0066 |
| | (-1.21) | (-1.28) | (-1.65) | (-0.81) | (-0.63) | (-1.22) | (-1.35) | (-1.28) |
| Return on Assets | 0.0193 | 0.0190 | 0.0206 | 0.0152 | 0.0219 | 0.0188 | 0.0182 | 0.0200 |
| | (1.08) | (1.06) | (1.19) | (0.88) | (1.06) | (1.06) | (1.02) | (1.13) |
| Log(Book-to-Market) | 0.0004 | 0.0005 | 0.0003 | 0.0003 | 0.0004 | 0.0005 | 0.0005 | 0.0004 |
| | (0.63) | (0.78) | (0.53) | (0.51) | (0.72) | (0.71) | (0.70) | (0.65) |
| Market Cap. | -0.0007* | -0.0007* | -0.0006* | -0.0008** | -0.0011*** | -0.0006 | -0.0007* | -0.0007** |
| | (-1.93) | (-1.92) | (-1.81) | (-2.15) | (-2.75) | (-1.62) | (-1.91) | (-1.99) |
| Observations | 492 | 492 | 393 | 252 | 492 | 492 | 492 | 492 |
| Adjusted $\%R^2$ | 20.21 | 20.24 | 21.27 | 21.97 | 20.24 | 20.25 | 20.24 | 19.78 |

**Table 9:** Earning Announcement Cumulative Abnormal Returns and Trade Size Clusters

I present Fama and MacBeth (1973) regressions based on the Cumulative Abnormal Return (CAR) surrounding the Earnings Announcement Dates (EAD). The analysis covers the period from days -1 to +1 relative to the EAD and includes equally-weighted results for NYSE/Amex/NYSE Arca firms between 1980 and 2021. To form distinct portfolios, trades are categorized based on established trade size clusters: 100 to 499 shares, 500 to 999 shares, 1000 to 4999 shares, 5000 to 9999 shares, and finally, trades of $\geq$ 10000 shares. I exclude the three days preceding the EAD when calculating the 15-day moving average (-3 to -18) to determine the percentage of trades within each trade size category. CAR's are determined using a five-factor Fama-French model. Additionally, firms with a share price of less than $5.00, measured three days before the EAD, are excluded from the analysis. I further divide the period into two sub-periods: 1980 to 2000 as the pre-decimalization period and 2001 to 2021 as the post-decimalization period. The results demonstrate statistical significance (*, **, ***) for the trade size clusters of 500 to 999 shares and 1000 to 4999 shares, with t-statistics shown in parentheses.

| Trade Group | 1980 to 2022 | 1980 to 2022 | 1980 to 2000 | 2001 to 2022 | 1980 to 2022 | 1980 to 2000 | 2001 to 2022 | 1980 to 2022 | 1980 to 2022 |
|---|---|---|---|---|---|---|---|---|---|
| 100 to 499 Shares | 0.0052 | | | | | | | | |
| | (1.55) | | | | | | | | |
| 500 to 999 Shares | | -0.0017*** | -0.0020** | -0.0015** | | | | | |
| | | (-3.13) | (-2.34) | (-2.08) | | | | | |
| 1000 to 4999 Shares | | | | | -0.0011** | -0.0012* | -0.0010* | | |
| | | | | | (-2.48) | (-1.82) | (-1.68) | | |
| 5000 to 9999 Shares | | | | | | | | 0.0003 | |
| | | | | | | | | (0.88) | |
| $\geq$ 10000 Shares | | | | | | | | | 0.0003 |
| | | | | | | | | | (0.82) |
| Intercept | 0.0014** | -0.0035** | -0.0028* | -0.0041 | -0.0023 | -0.0014 | -0.0031 | 0.0007 | 0.0005 |
| | (2.17) | (-2.19) | (-1.81) | (-1.55) | (-1.40) | (-1.18) | (-1.08) | (0.30) | (0.22) |