Do Individual Investors Ignore Transaction Costs?

Deniz Anginer, Snow (Xue) Han, Çelim Yıldızhan*

May 2025

Abstract

Using close to 800,000 transactions by 66,000 households in the United States and close to 2,000,000 transactions by 303,000 households in Finland, this paper shows that, on average, individual investors with longer holding periods choose to hold less liquid stocks in their portfolios. The relationship between holding periods and transaction costs is stronger among more financially sophisticated households. We confirm our findings by analyzing changes in investors' holding periods around exogenous shocks to stock liquidity. Our findings challenge the notion that individual investors ignore non-salient costs when making investment decisions and suggest that they are cognizant of the cost of trading stocks.

JEL Classifications: G11, G12, G14

Keywords: individual investors; liquidity; transaction costs; investor attention; behavioral bias

---

* Deniz Anginer is at Simon Fraser University and can be reached at deniz_anginer@sfu.ca. Snow (Xue) Han is at San Francisco State University and can be reached at hanx@sfsu.edu. Çelim Yıldızhan is at the Lee Business School at the University of Nevada, Las Vegas and at Koç University and can be reached at either celim.yildizhan@unlv.edu or cyildizhan@ku.edu.tr. The authors thank the Managing Editor of the Journal of Financial and Quantitative Analysis Hendrik Bessembinder as well as an anonymous referee at the same journal, Jussi Keppo, Jeff Netter, Paolo Pasquariello, Bradley Paye, Annette Poulsen, Nejat Seyhun, Tyler Shumway, Noah Stoffman, Uday Rajan, and seminar participants at the Ross School of Business at the University of Michigan, the Terry College of Business at the University of Georgia, the Erasmus Liquidity Conference, the Tel Aviv University Finance Conference, and the Western Finance Association meetings for comments and suggestions. We thank Terry Odean for providing us with the US household trading dataset and Jussi Keppo for providing us with the Finnish household trading dataset. All errors are our own. This paper is based on an earlier draft titled "Liquidity Clienteles."

1. Introduction

In theoretical models of liquidity, investors' expected holding periods determine how transaction costs are priced in asset values.[1] Long-term investors who can amortize trading costs over longer expected holding periods require lower per-period returns than investors with shorter expected holding periods. These models rely on the fundamental assumption that rational investors minimize per-period transaction costs on their investments. Counter to the idea in these theoretical models that investors understand and incorporate the impact of transaction costs in their investment decisions, findings in the behavioral finance literature suggest that individual investors tend to ignore non-salient costs when making investment decisions.

In this paper, we use trading records of households in the US and in Finland to investigate whether individual investors are cognizant of the costs of trading securities when making investment decisions. Specifically, we examine whether individual investors hold illiquid securities with high transaction costs longer as stipulated by theoretical models of liquidity pricing or ignore transaction costs as suggested by prevalent findings in the behavioral finance literature.

Existing evidence suggests that individual investors ignore non-salient costs as they relate to mutual fund fees. Barber, Odean, and Zheng (2005) show that individual investors pay attention only to the salient costs of mutual funds and ignore hidden operating costs. Consistent with these findings, Gil-Bazo and Ruiz-Verdu (2008, 2009) document a negative relationship between mutual funds' before-fee performance and the fees they charge investors. Surveys also suggest that retail investors do not understand all the costs associated with investing in mutual funds (Alexander, Jones, and Nigro 1998; NASD Investor Literacy Survey, 2003).[2]

[1] See for instance, Amihud and Mendelson (1986), Constantinides (1986), Vayanos (1998), Vayanos and Vila (1999), Heaton and Lucas (1996), Huang (2003), Lynch and Tan (2011), and Lo, Mamaysky, and Wang (2004).

[2] For example, only 21% of the retail investors that responded to the NASD Investor Literacy Survey (2003) knew the meaning of a "no load" mutual fund.

There is also evidence that individuals do not pay attention to non-salient costs in other domains. In a field experiment Hossain and Morgan (2006) show that buyers in eBay auctions ignore shipping costs when the price of the item being auctioned is much higher than the shipping costs. Chetty, Looney, and Kroft (2009) document that consumers underreact to taxes that are not salient. Similarly, Finkelstein (2009) finds that drivers are less aware of tolls paid electronically. These findings suggest that individual investors may not fully understand and incorporate non-salient transaction costs such as bid-ask spreads and price impact when trading.

Consistent with the notion that investors do not pay attention to non-salient costs, a number of studies have found that individual investors tend to overtrade and lose substantial amounts to transaction costs without any gain in performance. Barber and Odean (2000), for instance, show that while there is a minor difference between the gross performance of individual investors who trade frequently and those who trade infrequently, the net returns after transaction costs for infrequent traders are about 7% higher per year than those for frequent traders. Barber and Odean (2000) attribute their findings to individual investors' overconfidence. Barber et al. (2009) and French (2008) confirm this finding.[3]

However, losses incurred by individual investors after accounting for transaction costs would not necessarily imply that these investors are not paying attention to transaction costs. First, investors can trade for a variety of reasons other than information or behavioral biases, such as when they experience income shocks (Lynch and Tan 2011) or exogenous liquidity shocks (Huang 2003). Since we do not observe the full portfolios of individual investors, we cannot fully infer

[3] Using the complete transaction history of all investors in Taiwan, Barber et al. (2009) finds that individual investor losses due to transaction costs equal 2.2% of GDP, without any gain in performance. French (2008) finds that, each year, investors spend about 0.67% of the aggregate value of the market on transaction costs, again without any gain in performance. He estimates the capitalized cost of active investing to be at least 10% of the total market capitalization.

motivations behind their trades. Second, even if most of the overtrading by individual investors could be attributed to overconfidence, that would not necessarily imply that such investors do not pay attention to transaction costs.

In this paper, we directly test whether individual investors pay attention to transaction costs by examining the relationship between transaction costs and the holding periods of individual investors. Rather than focus on the trading performance of households, we analyze whether individual investors understand the trade-offs between holding periods and transaction costs. In doing so, our goal is not to offer an alternative setting to test the asset pricing implications of transaction costs. Rather, our focus in this paper is on the more specific question of whether and how retail investors incorporate transaction costs in their investment decisions.[4]

We model investors' holding periods as a function of transaction costs using close to 800,000 transactions made by 66,000 households in the US, and 2,000,000 transactions made by 303,000 households in Finland. We use survival analyses and model investors' sell versus hold decisions at each point in time as a function of transaction costs using hazard regressions.

We find that transaction costs are an important determinant of investors' holding periods after controlling for various household and stock characteristics. We find that in the US a stock in the highest transaction cost decile (quintile) is 40% (20%) less likely to be sold than a stock that has lower transaction costs but with similar firm and investor characteristics, consistent with the predictions of theoretical models of liquidity.

[4] A number of papers examine the pricing impact of holding periods as measured by turnover on stock returns (Atkins and Dyl 1997; Datar, Naik, and Radcliffe 1998; Hu 1997). While examining the pricing impact of average turnover is informative from an asset pricing perspective, it does not tell us how individual investors incorporate transaction costs in their investment decisions. Examining market averages can also mask large cross-sectional variation and skewness in the holding periods of investors for the same stock. For instance, some stocks can be more heavily traded by institutional investors, or market makers could be more active in some stocks than others. Some stocks may thus have a group of exceedingly long holding period owners, but high turnover among the smaller group of remaining investors.

We check the validity of this finding by replicating our analyses using an alternative transactions dataset from Finland, which serves as an "out-of-sample" verification. Almost identical to the US results, we find that an otherwise similar stock in the highest transaction cost decile in Finland is 40% less likely to be sold compared to a stock that has lower transaction costs. Since the data from Finland include the complete transactions of all Finnish households between 1995 and 2003, the results suggest that our findings can be generalized to the full cross-section of households. Our results remain robust to controlling for firm- and household-specific effects, additional controls, and alternative measures of transaction costs.

We also find that households differ in how much attention they pay to the transaction costs of the securities they trade. We find that investors who are more financially sophisticated pay more attention to transaction costs. We follow Goetzmann and Kumar (2008) and assume that financial sophistication is correlated with education, occupation, and monetary resources available to an investor. We also use information contained in investors' trades to identify sophisticated investors. We classify households that have above average income, hold technical and managerial positions, trade options, invest in foreign securities, have short positions, and have a portfolio that is more concentrated than the median portfolio concentration as financially more sophisticated. Our findings suggest that investor sophistication plays a role in how much attention investors pay to transaction costs. We confirm our findings on financial sophistication using data from Finland.

There is likely to be endogeneity in the relationship between holding periods and measures of transaction costs used in this paper. For instance, as trading interest in a stock increases, the costs associated with trading that stock decrease.[5] In order to address potential endogeneity concerns, we study investor behavior around two quasi-exogenous liquidity shocks.

First, we examine how holding periods change around stock split events. An extensive line of literature documents significant reductions in transaction costs and an increase in liquidity after stock splits.[6] Consistent with the prior literature, we first verify that transaction costs decrease (stock liquidity increases) subsequent to stock splits in our sample period. We then show that investors' average holding periods decline in response to the increase in liquidity following stock splits. Our results suggest that the probability of sale by an average investor increases by 16% in the 6-month time period after a stock split.

Second, we conduct an event study around the reduction in the minimum tick size for stocks priced between one and five dollars listed on the American Stock Exchange. On September 3, 1992, the American Stock Exchange (AMEX) reduced its minimum price increment from 1/8th of a dollar to 1/16th of a dollar for stocks priced between \$1 and \$5. One of the motivations for this change was to reduce bid-ask spreads. Several papers document that both quoted and effective spreads declined after this change, leading to lower transaction costs (Ahn, Cao, and Choe 1996; Crack 1996).

We investigate the holding period decisions of investors for stocks impacted by the tick-size change. Specifically, we compare the differential impact of the rule change on the holding period decisions of investors in the treated firms (AMEX stocks priced \$1 to \$5) versus in three alternative

[5] We should note, however, that the baseline or the average transaction cost of a given stock is likely to change slowly over time and is likely to be stable during short time periods in the absence of corporate events. For instance, the liquidity level of a penny stock would increase with increased trading interest, but it is not likely to achieve the same level of liquidity of a large cap stock purely based on investor interest or attention.

[6] For example, Schultz (2000) shows that the number of trades, especially small trades, increases significantly after stock splits. Desai, Nimalendran, and Venkataraman (1998) find that both informed trades and noise trades increase after stock splits. Kryzanowski and Zhang (1996) show that absolute trading volumes of Canadian stocks increase subsequent to stock splits. Conroy, Harris, and Benet (1990) also show a significant reduction in the absolute bid-ask spread following stock splits.

groups of control firms. The first control group contains all firms on AMEX that were priced at or above $5 at the time of the rule change. The second control group contains firms that were priced between $1 and $5 but listed on the NYSE and NASDAQ exchanges and as a result were not affected by the tick size change. The third group contains all non-affected stocks on the three major exchanges. In all three comparisons, we find that the tick size reduction led retail investors to reduce their holding periods in the treated firms in reaction to reduced transaction costs. We find that investors' likelihood of selling their impacted shares significantly increased around the tick size reduction. On average, investors were 16.7% more likely to sell an impacted stock (AMEX stock priced $1 to $5) in the six months subsequent to the tick size change rule, controlling for stock characteristics.

The remainder of this paper is organized as follows. Section 2 develops the hypotheses we evaluate in the paper. Section 3 describes the individual transactions datasets and the construction of the main variables used in this study. Section 4 reports our main results about the relationship between transaction costs and holding periods. Section 5 provides robustness tests to address concerns that holding periods are determined endogenously and also uses individual transactions from Finland as an out-of-sample test to verify the US results. Section 6 concludes.

2. Hypotheses

A number of theoretical models link the level of transaction costs to expected holding periods of investors. In a seminal paper, Amihud and Mendelson (1986) develop a model in which investors with different exogenous holding periods trade securities with fixed transaction costs. They show that transaction costs result in a clientele effect where investors with longer holding periods choose to hold illiquid stocks in equilibrium. This equilibrium results from rational investors trying to minimize amortized transaction costs over their holding periods. In the model, the expected gross return becomes an increasing and concave function of relative transaction costs. Amihud and Mendelson find empirical support for this hypothesis using spreads and stock returns over the 1961 to 1980 time period.

While Amihud and Mendelson's (1986) model assumes that the holding periods of investors are exogenously determined, later studies have extended this model to incorporate dynamic decisions of investors and make holding periods endogenously determined. In models where the marginal utility from trading is low (Constantinides 1986; Heaton and Lucas 1996; Vayanos 1998; Vayanos and Vila 1999), investors respond to transaction costs by turning over their portfolio less frequently. These models predict a liquidity premium on asset prices that is significantly lower than transaction costs, but they also predict unrealistically low levels of trading volumes as investors respond to higher transaction costs by lowering their trading activity. In models where investors trade more frequently (Huang 2003; Lynch and Tan 2011; Lo, Mamaysky, and Wang 2004) the resulting liquidity premium can be large. While these dynamic models differ in their assessments regarding how transaction costs are priced, they share a common assumption that holding horizons are the outcome of optimal investor behavior, and that investors rationally trade off the costs and benefits of delaying trades. As theoretical models predict that households' holding periods across various assets in their portfolios are positively related to transaction costs, our first hypothesis is:

H1: *Holding periods of households across stocks are positively related to measures of transaction costs after controlling for investor and stock characteristics.*

Previous studies have shown that, on average, households' stock investments perform poorly. Odean (1999), for instance, reports that individual investors' purchases under-perform their sales by a significant margin. However, other studies have shown that there exists a subset of retail investors who display greater financial sophistication and market understanding than the average retail investor. For example, Coval, Hirshleifer, and Shumway (2005) document strong persistence in the performance of individual investors' trades and show that some skillful individual investors can earn positive abnormal profits across different periods. Ivkovic, Sialm, and Weisbenner (2008) propose and empirically document that individual investors who hold more concentrated portfolios have better stock-picking skills that allow them to outperform other investors. Feng and Seasholes (2005) find that investors who are more sophisticated and possess more trading experience suffer less from the disposition effect bias.

Given that previous studies have documented heterogeneity in the performance and investment decisions of individual investors, we expect to find similar cross-sectional differences in the correlation between holding periods and transaction costs among households. In particular, we expect that individual investors who are more financially sophisticated make better decisions and pay closer attention to transaction costs. We follow the extant literature and assume that financial sophistication is correlated with education, occupation, and monetary resources available to an investor. We also use information contained in investors' trades to identify sophisticated investors. Our second hypothesis is:

H2: *The correlation between holding periods and transactions costs is stronger for financially more sophisticated investors.*

3. Data

3.1 Household Transactions and Demographics Information

This study uses two datasets to analyze the trading behaviors of households. The first dataset contains transactions for a subset of individual investors in the United States, while the second contains transactions of all investors in Finland. The individual trade data for the United States come from a major US discount brokerage house that recorded the daily trades of 78,000 households from January 1991 to December 1996. This is the same dataset as used in Barber and Odean (2000).[7] We focus only on the common stock transactions of households in this study, which account for nearly two-thirds of the total value of household investments. We exclude from the current analysis investments in mutual funds, American Depositary Receipts (ADRs), warrants, and options.

Our final sample includes over 66,000 households with close to 800,000 transactions. The dataset includes for each transaction, the number of shares traded, the transaction price, and value of the position at market close. The dataset also includes demographic information for a smaller subsample of households (37,664 households), such as income, age, gender, occupation, and marital status.

To address concerns that our findings may be specific to the data and sample period we study, we repeat our analyses using an alternative transaction dataset from Finland. This dataset comes from the central register in the Finnish Central Securities Depository (FCSD). The register officially records all the trades of all Finnish investors (both individual and institutional) daily from January 1995 to December 2003. Compared to the US dataset, the Finnish dataset has better coverage as it includes the complete trading records of all market participants rather than a subset of market participants. For the purposes of this study, we ignore institutional trades and utilize only the trades of individual investors in Finland. Like the US dataset, the Finnish dataset reports each transaction, the number of shares traded, the trading price, and the daily closing price. We

[7] For a more detailed description of this dataset, please refer to Barber and Odean (2000, 2001). A comparison of this dataset with Survey of Consumer Finances, IRS and TAQ data has shown it to be representative of US individual investors (Barber, Odean, and Zhu 2006; Ivkovic, Sialm, and Weisbenner 2008; Ivkovic, Poterba, and Weisbenner 2005).

can also observe the initial holdings for each account at the beginning of the sample period, which allows us to keep track of the daily holdings of households. While the dataset reports demographic information such as age and gender for a subset of investors, it does not include information about income, occupation, and marital status. A more detailed description of the Finnish dataset can be found in Grinblatt and Keloharju (2000, 2001). To calculate stock and firm characteristics for the Finnish stocks, we obtain data from Datastream.

We use the following investor characteristics in our analyses: investor age (*Age*), log of annual income in dollars (*Log (Income)*); a dummy variable that is equal to one if the trader is married (*Married Dummy)*; a gender dummy that is equal to one if the trader is male (*Male Dummy)*, a dummy to capture if the trader holds a technical or managerial position (*Professional Dummy)*; a dummy that takes on the value of one if the trading account is a retirement account (*Retirement Acct Dummy)*; and a dummy that equals one if the trader is retired (*Retired Dummy)*.

We also identify certain trader characteristics from each household's trading history and define the following control variables: *Foreign Securities Dummy*, equals one if the household has ever traded foreign securities; *Option User Dummy*, equals one if the household has ever traded options; and *Short User Dummy*, equals one if the household has ever held a short position. We also estimate the log of the average total dollar value of each household's equity investments, *Log* (*Equity Portfolio Value*). Finally, we estimate the concentration of each household's portfolio (*Portfolio Concentration*) computed as the sum of the squared value weights of each stock in a household's portfolio following Ivkovic et al. (2008). We calculate the value of equity investments monthly and calculate all the other trade and investor characteristics over the full sample period.

3.2 Measures of Transaction Costs and Firm Controls

Transaction costs are multifaceted and are usually defined in terms of the costs and risks associated with trading financial securities. These costs incorporate price impact, asymmetric information, and inventory risk. A number of different measures of transaction costs have been proposed and used in the literature. Instead of relying on a single measure, we use several different measures that are commonly used in previous papers and can be estimated for both the U.S and Finland datasets.

The first measure is the Amihud illiquidity ratio ($Illiq$) from Amihud (2002), calculated as:

$$Illiq_{i,t} = \frac{1}{D_{i,t}} \sum_{d=1}^{D_{i,t}} \frac{|r_{i,d}|}{dvol_{i,d}} \quad (1)$$

where $r_{i,d}$ is the daily return for stock *i* in day *d*. $dvol_{i,d}$ is the dollar volume for stock *i* in day *d*. $D_{i,t}$ is the number of trading days in month *t*. The Amihud measure is similar to Kyle's lambda and captures the price impact of trades over a specific time period. Following Acharya and Pedersen (2005), we adjust the Amihud measure as in the following to remove outliers and to make it stationary: $AdjIlliq_{i,t} = \min\,[0.25 + 30 \times Illiq_{i,t} \times M_{t-1}, 30]$, where M_{t-1} is the ratio of the capitalizations of the market portfolio at the end of the month *t-1* to that of the market portfolio in July 1962. The higher the adjusted Amihud ratio, the more illiquid the stock is.

The second measure uses the proportion of trading days with zero returns (*Zerofreq*) to capture transaction costs. Following Lesmond, Ogden, and Trzcinka (1999), we compute the proportion of days with zero returns for each stock each year as *Zerofreq*. The higher the *Zerofreq*, the more illiquid the stock is.

We also compute a number of measures using intra-day trades for the US sample. We use a 5-second delay to match trades with quotes and apply the same filters discussed in Hvidkjaer (2006).

Effective Spread / Price is the difference between the transaction price and the quoted bid-ask midpoint multiplied by two and scaled by transaction price. *Relative Spread / Price* is defined as the quoted bid-ask spread divided by transaction price, and *Relative Spread / Mid* is defined as quoted spread scaled by the bid-ask midpoint. *Depth* is defined as the midpoint of bid size and offer size (both in number of round lots). As depth tends to be skewed, we use *log(1+depth)* in our analyses. To reduce potential endogeneity arising from contemporaneous measurement and to smooth out idiosyncratic changes, we use the 12-month moving average of each liquidity measure in our analyses. The liquidity measures are calculated at the end of the month prior to transactions. For instance, if an individual investor has a sell transaction on May 15, the liquidity measures would be calculated as of April 30.

Finally, we use the actual trades of investors to measure realized transaction costs following Barber and Odean (2000). We estimate closing price spread (*Closing price Spread %*) for purchases as the negative of the closing price from CRSP divided by the transaction price minus one. Closing price spread for sales is equal to the closing price from CRSP divided by the transaction price minus one. We also calculate commissions (*Commission %*) as the amount charged by the brokerage for the trade scaled by the dollar value of the trade. In the analyses, we use the sum of commissions and purchase spread (*Closing price Spread % + Commission %*).

We control for a number of firm characteristics in the analyses. These are firm size measured by log of market capitalization (*Size*), book-to-market ratio (*B/M*), momentum calculated using returns over the past 12 months excluding the previous month (*Momentum*), idiosyncratic volatility (*Ivol*), maximum daily return over the past one month (*MaxPrc*), and the CAPM Beta (*Beta*). We also control for *Unrealized Returns* as (selling price - purchase price) / purchase price to capture potential disposition bias. In calculating *Unrealized Returns*, if a sale is never observed and the sale price is unavailable, we use the stock price on the last day of our sample period. With the exception of unrealized returns which are calculated on the transaction date, all other stock characteristics are calculated as of the end of the month prior to the transaction. For instance, if an individual investor has a sell transaction on May 15, the stock characteristics would be calculated as of April 30.

Table 1 reports the summary statistics for stock and investor characteristics for the US. Panel A reports descriptive statistics for stocks that are traded by households in the dataset. For comparison, panel B provides descriptive statistics for the CRSP stock universe during the same sample period. Summary statistics are calculated by pooling annual stock-level observations from 1991 to 1996. Panel A and B show that the price, size, book-to-market ratio, and past returns for stocks in our sample are similar in magnitude to those in the entire CRSP universe. For example, the median, 25th percentile and 75th percentile prices are the same for our sample of stocks and for those in the CRSP universe. The average book-to-market ratio for our sample of firms is 0.78, which is slightly higher than the average book-to-market ratio of 0.72 for the CRSP universe, while the median is 0.57 for our sample and 0.56 for the CRSP universe. The average and median sizes of our sample firms are also slightly larger than those of the CRSP universe. The differences between the two samples and their statistical significance are reported in the last column in Panel B. The transaction costs measures are marginally lower in our sample compared to the larger CRSP universe. Overall, the differences are economically small, indicating that our sample of stocks are representative of the entire stock market during the sample period.

Panel C reports the summary statistics for the US individual investor characteristics. The majority of the investors are in their 40s and 50s, with an average (median) age of 49.58 (48), and 15% of the investors are retired. Only 10% of the primary US account holders for the transactions analyzed in this study are female, 76% are married, and 66% hold technical or managerial positions. The mean (median) portfolio value is $80,342 ($22,952) for the households analyzed in this study, and the mean (median) annual income is $76,840 ($87,500) for these investors over the sample period. In addition, 14% of the households have traded options, 22% have traded foreign securities, and 38% have held a short position at some time over the sample period analyzed. The mean (median) US individual investor's portfolio concentration is .52 (.48), which roughly corresponds to holding two stocks with equal weights.

4. Transaction Costs and Holding Periods in the US

4.1 Holding Periods and Transaction Costs

In this section, we provide empirical evidence in support of the first hypothesis (*H1*). We begin by computing a holding period for each transaction in the dataset. The holding period for a transaction is defined as the number of trading days from the first purchase to the first sale of that stock, following the approach of Seru, Shumway, and Stoffman (2010). This generates 799,469 holding period observations, with a median (mean) of 207 (550) trading days for retail investors in the United States.

We begin our analyses by sorting stocks into two broad transaction cost groups each year based on our main transaction cost measure, namely the adjusted Amihud illiquidity measure. One group consists of stocks in the highest transaction cost decile while the other group comprises the rest of the stocks in the other nine deciles. We plot Kaplan-Meier survival probabilities for these two broadly defined groups of stocks in Figure 1. The x-axis shows the number of days that have passed since the purchase of a representative stock in each group, while the y-axis represents the probability that the investor will continue to hold this representative stock conditional upon no sale up to that point in time. The solid line plots the survival probability of a representative stock in the highest transaction cost decile, while the dashed line graphs the survival probability of a representative stock for the other nine deciles. In Figure 1a, we plot survival probabilities for stocks in the US, while in Figure 1b, we plot survival probabilities for stocks in Finland. Investors are more likely to sell stocks with lower transaction costs as the survival probabilities are lower for these stocks. The figures provide preliminary evidence that holding periods are strongly related to measures of transaction costs.[8]

We use a hazard model to analyze the relationship between holding periods and transaction costs controlling for the confounding effects of stock and investor characteristics.[9] Specifically, we model investors' sell versus hold decision using a Cox proportional hazard model with time-varying as well as static explanatory variables. The hazard model takes the following form:

$$h(t) = h_0(t) exp(\beta' X + \theta' Z_t) \quad (2)$$

The left-hand side variable, $h(t)$, is the hazard rate, the probability of selling a stock on day t conditional upon holding that stock until that point (t) in time. X is a vector of explanatory variables which are static and do not change over time. Z_t represents a vector of time-varying covariates which can take on different values at different points in time.[10] $h_0(t)$ is the baseline hazard rate and describes the hazard rate when the independent covariates are all equal to zero. Using the Cox (1972) estimator, we can estimate coefficients on X and Z_t without specifying a baseline $h_0(t)$

[8] In Table A1 in the Online Appendix (Anginer, Han and Yıldızhan, SSRN), we show that a stock in the highest quintile illiquidity group is approximately 0.8 times as likely (20% less likely) to be sold as a stock not belonging to that group.

[9] The hazard model framework has been used in the past by Seru, Shumway, and Stoffman (2010) as well as Feng and Seasholes (2005) to model holding periods of individual investors.

[10] The static variables are the demographic variables (*Age, Log (Income), Married Dummy, Male Dummy, Professional Dummy, Retirement Acct Dummy, Retired Dummy*), and most of the trade variables (*Foreign Securities Dummy, Option User Dummy, Short User Dummy, Portfolio Concentration*). The variables that vary over time are stock characteristics (*Size, B/M, Momentum, Beta, Ivol, MaxPrc, Unrealized return*), and the natural logarithm of the monthly equity value of the investors (*Log (Equity Portfolio Value)*).

hazard rate. Positions that are not closed by the end of the sample period are treated as censored observations.

We control for investor characteristics that are directly observable such as age, income, gender, marital status, employment status, and occupation, as well as another set of less readily observable variables that are extracted from investors' positions and trades, such as the total wealth invested in their portfolios and whether the individual investors ever short stocks, trade options, or trade foreign securities. We also control for size, book-to-market ratio, and momentum to account for investors' preferences for stocks with certain characteristics that are known to be associated with expected returns.

As there is also likely to be seasonality in purchases and sales, we further include calendar year and month dummies in the hazard regressions. Open stock positions, for instance, may be closed out in December for tax reasons. Finally, we use unrealized gains/losses as a control variable. Although momentum does capture the effect of past returns on trading decisions, unrealized gains and losses for each individual investor could be different based on the original purchase price.

Table 2 reports the results from the hazard regressions. Following standard reporting conventions, we report the hazard ratios instead of the estimated coefficients. The hazard ratio is similar to the odds ratio estimated from a binary choice model and is defined as the ratio of two hazard rates when one explanatory variable is changed by one unit from zero holding all other variables constant. A hazard ratio of less than one would suggest that the explanatory variable reduces the probability of selling the stock. In contrast, a hazard ratio larger than one would suggest that a higher exposure to the explanatory variable would increase the likelihood of selling the stock, thus reducing the likelihood that the investor would continue holding on to the stock.

In Table 2, we report the results using only the adjusted Amihud illiquidity ratio for all specifications. We provide results for alternative measures of transaction costs in Table 3. Column (1) of Table 2 shows that the estimated hazard ratio for the adjusted Amihud illiquidity ratio is 0.981 when we do not control for stock or investor characteristics. It is less than one and statistically significant, suggesting that the sale probability of a stock declines with higher transaction costs. Specifically, the average investor would be 9.3% less likely to sell a stock in the 75^{th} percentile in terms of illiquidity compared to a stock with a median level of illiquidity using the adjusted Amihud illiquidity ratios.[11]

As households could have different preferences and potentially have different holding periods, we control for heterogeneity across households within the hazard framework. We assume different baseline hazard rates for each household and estimate a model with partial likelihood stratification. The household level stratification allows for the possibility of each household having a different baseline holding period, which is analogous to using household fixed effects in OLS regressions. Similarly, we use firm stratification to allow for the possibility that each stock has a different average holding period. In column (2) of Table 2, we calculate hazard ratios using firm and household stratifications to account for household and firm fixed effects. The estimated hazard ratio for the adjusted Amihud measure (*AdjIlliq*) is 0.973 and statistically significant, consistent with earlier results. Controlling for household and firm level fixed effects suggests that a one standard deviation increase in the adjusted Amihud illiquidity ratio would reduce the sale likelihood by 18.5%.

[11] The median adjusted Amihud illiquidity ratio is 1.18 and its 75^{th} percentile is 6.26 for our sample stocks. Moving from the median stock to the 75^{th} percentile stock would result in an increase of 5.08 in the adjusted Amihud illiquidity ratio. As the hazard ratio for the adjusted Amihud ratio (*Adjilliq*) is 0.981, an investor would be exp(ln(0.981)*5.08) = 0.907 as likely to sell the stock in the 75^{th} percentile of adjusted Amihud illiquidity as a stock with median adjusted Amihud illiquidity. This difference in illiquidity reduces the likelihood of sale by 9.3% (=1-0.907).

Controlling for heterogeneity among households and stocks leads to stronger results as the hazard ratio is reduced from 0.981 to 0.973. To better understand the source of this variation, we run a regression of holding periods on household and stock fixed effects. We find that household fixed effects explain about 35% of the cross-sectional variation in holding periods, while stock fixed effects explain about 18% of the variation. These results suggest that both household and stock fixed effects influence holding periods and that households also differ in their baseline holding periods.

We examine in detail how specific stock and investor characteristics affect households' trading decisions. We add stock characteristics first in column (3) of Table 2, and then further control for investor characteristics and unrealized returns in column (4). Since demographic information is only available for a subset of investors in the dataset, the number of observations reported is lower in column (4). Our initial finding on transaction costs is unchanged with these additional controls. The loading on the *AdjIlliq* in column (3) is still less than one at 0.981 and statistically significant. The estimated hazard ratio for momentum is statistically significant and larger than one (1.135), which indicates that investors are more likely to sell recent winners. More specifically, a one standard deviation increase in the past 10-month momentum returns (from month-12 to month-2) would increase the probability of sale by 30.6%. The estimated hazard ratio for size is 0.649 and that for the book-to-market ratio is 0.681, both of which are less than one and are economically and statistically significant, suggesting that US individual investors tend to hold large and value stocks for longer periods.

In column (4) we control for *Unrealized Returns*, to account for the impact of disposition effect - the tendency of individual investors to hold on to losing stocks for too long and to sell winners too quickly on our results. Our basic inferences regarding the impact of transaction costs on retail investors' holding periods are unaffected when we control for unrealized returns along with trader demographics and trade characteristics. The estimated hazard ratio on *AdjIlliq* is 0.975, comparable to our findings in the initial three columns. The coefficient on *Unrealized Returns* is statistically significant and greater than one (1.134), suggesting that retail investors are more likely to sell shares that have higher unrealized returns. This finding is consistent with the disposition bias documented in the literature.

For robustness, we repeat our main analyses by excluding holding periods less than two days to remove day traders. Specifically, in column (5) we repeat our analyses in column (4) by excluding observations with one-day and two-day holding periods. Our main findings are unaffected by this restriction.[12]

For robustness, we also control for additional variables that prior studies have shown to affect individual investor trading decisions. Prior studies have shown that individual traders tend to buy attention-grabbing stocks.[13] To control for investor attention, we add stock characteristics that are positively correlated with investor attention to our baseline hazard regression. The stock characteristics we use are idiosyncratic volatility (*Ivol*), maximum daily return over the past one month (*MaxPrc*), and CAPM Beta (*Beta*). The hazard regression results with these controls are reported in column (6) of Table 2. *Beta* and *Ivol* have statistically significant hazard ratios of greater than one, 1.111 and 2.807, respectively, while the estimated hazard ratio for *MaxPrc* is almost one and statistically insignificant. These results are consistent with investors trading

[12] In Online Appendix (Anginer, Han and Yıldızhan, SSRN) A.2, we repeat the analyses for all specifications, removing observations with one-day and two-day holding periods.

[13] Barber and Odean (2008) document that individual investors tend to buy attention-grabbing stocks, such as stocks with extreme one-day returns, which is also supported by Bali, Cakici, and Whitelaw (2011). Bali, Engle, and Tang (2017) show that stocks with high conditional betas are also attention-grabbing and attract individual investors. Kumar (2009) shows that individual investors prefer lottery-like stocks.

attention-grabbing stocks more frequently. The estimated hazard ratio in column (6) for *AdjIlliq* remains significant at 0.977 and is similar in magnitude to the hazard ratio reported in column (4).

We conduct an alternative analysis in Table A1 in the Online Appendix (Anginer, Han and Yıldızhan, SSRN) section A.1. Specifically, we rank all stocks by the Amihud illiquidity ratio and create a dummy variable (*AdjIlliq Dum*) that takes on a value of one if stock belongs to the highest illiquidity quintile. The use of the dummy variable makes it easier to interpret our results in Table 2. We find in Table A1 that in the US a stock in the highest transaction cost quintile is 20% less likely to be sold than a stock that has lower transaction costs but with similar firm and investor characteristics, consistent with other results in the paper as well as with the predictions of theoretical models of liquidity.

We repeat our analyses using six alternative measures of illiquidity described earlier, namely, *Zerofreq*, *Closing price Spread + Commission (%)*, *Effective Spread/Price (%)*, *Relative Spread/Price (%)*, *Relative Spread/Mid (%)* and *Log (1+depth).* In Panel A of Table 3, we repeat our analysis conducted in column (1) of Table 2 using these alternative measures. All estimated hazard ratios in columns (1) through (6) are less than one and statistically significant. Our results are qualitatively similar regardless of the illiquidity measure we use. The economic significance levels of these variables are also similar to those using the adjusted Amihud measure. For example, the estimated hazard ratio for *Closing price Spread + Commissions (%)* reported in column (2) is 0.945. This suggests that a one standard deviation increase in *Closing price Spread + Commissions (%)*, would lead to a 15.37% reduction in the average household's likelihood to sell.

It is possible that individual investors may care more about the trading costs incurred at the time of purchase rather than at the time of sale. To better understand whether there is an asymmetry in how purchase and sale transaction costs incurred are incorporated in holding period decisions, we investigate the impact of buy and sell transactions separately. Following Barber and Odean (2000), we calculate the closing price spread for purchases and sales separately for each trade. These results are reported in Table A3 in the Online Appendix (Anginer, Han and Yıldızhan, SSRN) section A.3. We repeat the analyses for Finland and report the relevant results in Table A4 in the Online Appendix (Anginer, Han and Yıldızhan, SSRN) section A.4. The coefficients on the purchase and sale transaction costs are similar in magnitude. Overall, these results are consistent with investors incorporating transaction costs incurred both at the time of purchase and at the time of sale.

4.2 Investor Sophistication

In this section, we investigate the impact of heterogeneity across households on the relationship between transaction costs and holding periods of investors. We provide empirical evidence in support of our second hypothesis (*H2*).

Following Goetzmann and Kumar (2008), we assume that financial sophistication is correlated with education and resources available to each investor. We create a sophistication measure based on household and trade characteristics. Specifically, we use seven criteria to construct our sophistication measure (*Sophistication*), which increases by one with each of the seven criteria being met. The criteria include: if the investor has income greater than $75K; if the investor works in a technical or managerial position (*Professional Dummy* = 1); if the investor is ranked among the top 25% of all investors in terms of total equity holdings; if the investor has ever traded an option (*Option User Dummy* = 1); if the investor has ever traded in foreign securities (*Foreign Securities Dummy* = 1); if the investor has ever shorted any equity (*Short User Dummy* =1); and if the investor's portfolio concentration is greater than 0.48, the median investor's level of portfolio concentration. The last criterion is based on findings in Ivkovic, Sialm, and Weisbenner (2008), who propose and empirically document that investors who hold more concentrated portfolios are financially more sophisticated as they possess informational advantages that allow them to outperform investors with diversified portfolios. The value of *Sophistication* ranges from a minimum of 0 for the least sophisticated investors to a maximum of 7 for the most sophisticated investors.

We sort investors into three groups based on their sophistication scores. Group 1 includes those with sophistication scores between 0 and 2, Group 2 is for investors whose sophistication scores are between 3 and 5, and Group 3 contains the most sophisticated investors with scores of 6 or 7. We then run a separate hazard regression for each of these three sophistication groups and examine how the relationships between transaction costs and holding periods change among investors with different levels of sophistication. Since many of the demographic variables and trade characteristics are used to calculate the sophistication score, these variables are not included as independent variables in our analyses in Table 4. Columns (1) to (3) of Table 4 report the estimated results for the least sophisticated group of households, the medium sophistication group, and the most sophisticated group, respectively.

We find that the coefficient on the adjusted Amihud illiquidity measure, *AdjIlliq*, is significantly positive and less than one for all sophistication groups. The estimated hazard ratios decrease monotonically from 0.984 for Group 1 to 0.975 for Group 2, and to 0.948 for Group 3 (the most sophisticated households). The estimated hazard ratio for *AdjIlliq* is 0.984 in column (1), indicating that an investor in the least sophisticated group would be 0.921 as likely to sell the stock in the 75th percentile of *AdjIlliq* as a stock with a median *AdjIlliq*. This would make them 7.9% less likely to sell. Similarly, the estimated hazard ratios for *AdjIlliq* in columns (2) and (3) would suggest that retail investors in Group 2 would be 0.879 as likely and those in Group 3 would be

0.762 as likely to sell the stock in the 75th percentile of *AdjIlliq* as a stock with median *AdjIlliq*. In other words, retail investors in the medium sophistication group would be 12.1% less likely to sell, while those in the most sophisticated group would be 23.8% less likely to sell when transaction costs increase. Overall, these results are consistent with our second hypothesis (*H2*) that financially more sophisticated investors pay closer attention to the impact of transaction costs when they trade.

5. Robustness

In this section, we conduct three additional analyses to show that our results are robust to potential endogeneity and selection concerns. If our transaction cost measures are related to certain unobserved variables which affect holding periods, then our results could suffer from an omitted-variables problem. To address this concern, we study two quasi-exogenous shocks to transaction costs. First, we use stock split events as quasi-exogenous shocks to transaction costs and examine investors' holding period decisions around stock split events in section 5.1. Second, we conduct an event study around the American Stock Exchange reduction of the minimum tick size from 1/8th of a dollar to 1/16th of a dollar for stocks priced between one to five dollars in 1992 and investigate the impact of this change on investors' holding period decisions and report our findings in Section 5.2. The third robustness test is meant to address potential selection issues with the US sample. The transaction-level dataset used in the US captures only a fraction of the US households' trades during certain years and hence may be insufficient to evaluate our main predictions. To address this criticism, we repeat our main analyses in section 5.3 utilizing another dataset which covers individual investors' complete trading records in Finland. Using an additional dataset from another country provides us with an "out-of-sample" test of our main findings.

5.1 Stock Splits

An extensive literature documents a significant reduction in transaction costs and improved liquidity subsequent to stock splits (Conroy, Harris, and Benet 1990; Desai, Nimalendran, and Venkataraman 1998; Kryzanowski and Zhang 1996; Schultz 2000). There is also evidence of a positive abnormal return reaction on the split announcement day for splitting firms and findings that post-split performances of splitting firms are statistically indistinguishable from those of similar non-splitting firms in the long-run (see for instance, Byun and Rozeff 2003).

We first verify empirically that stock splits indeed increase liquidity and reduce transaction costs. We identify a total of 3,586 stock splits that took place in the US between 1991 and 1996 for our sample. We remove reverse splits and splits that have a split factor of less than 0.25 (717 in total). Our final sample includes 2,869 forward split events.[14] Consistent with the findings in the literature, in Table A5 in the Online Appendix (Anginer, Han and Yıldızhan, SSRN), we show that there is a significant decline in our main measure of transaction costs (*AdjIlliq*) after a split event for stocks in our sample.

If investors hold illiquid securities for longer periods, then the reduction in transaction costs after stock splits should lead to shorter holding periods. We examine individuals' trading behavior over the same 6-, 9- and 12-month periods after a split event using a dynamic hazard regression framework. To construct the appropriate dataset for the dynamic hazard regression, we split the duration of a position into multiple periods.[15] The first period covers the time period before the split event. In this first period (pre-event), we assign a value of zero to the *After-Split Dummy*. The

[14] For robustness, we repeat our analyses by further removing 1,019 forward splits that coincide with the distribution of cash dividends within a [-30, +30] days window around the split event. When we use the remaining 1,850 "pure" forward splits, we obtain results that are qualitatively and quantitatively similar.

[15] Our approach follows that of the seminal paper titled "Mortality after the Hospitalization of a Spouse" by Christakis and Allison (2006).

second part is the time period from the split until the end of the event window of interest (i.e., three windows with a length of 6, 9, and 12 months). For the second period, *After-Split Dummy* takes on a value of one. The third period corresponds to the time-period after the split window (post-event), during which the *After-Split Dummy* takes on a value of zero.[16] For non-splitting stocks, the *After-Split Dummy is always zero.* In the analyses of forward splits, we exclude reverse splits.

Since it is possible for transactions to be open 6, 9, or 12 months after a split, this setup ensures that *After-Split Dummy* will only equal one when a sale event falls within the event window, and as time elapses to the post-event window period, *After-Split Dummy* will switch back to 0. *After-Split Dummy* captures the marginal impact of stock splits on sale decisions over a distinct event horizon. Since the baseline hazard rate in the Cox regression model captures the increasing probability of a sale as time passes, *After-Split Dummy* captures the marginal impact of being in the split window period on the probability of a sale and does not simply capture a mechanical relationship due to the fact that probability of a sale increases as time passes on.

Table 5 reports the estimated results of dynamic hazard regressions. All the regression models control for stock characteristics: size, book-to-market, and momentum, as well as calendar year and month effects. It is possible that split-event returns may lead to second-order effects that may influence investors' trading decisions. To control for the impact of post-split returns, we calculate split-event returns for each period and control for these returns in models (2), (3), (5), (6), (8), and (9) in the table. Finally, we account for the possibility that stock splits may lead to clientele effects: forward (reverse) splits may attract clienteles that prefer lower-priced (higher-priced) equities. Columns (3), (6), and (9) address the clientele issue by controlling for stock prices at the time of sale. If no sale takes place until the end of the dataset, then we use the last observed stock price.

[16] In the rare instances where there are multiple splits before a transaction is closed, the After-Split dummy will be one during the post-split window but will switch back to zero after each post-split event window.

Given the reduction in transaction costs after stock splits, we expect households to be more likely to reduce their holding periods, and thus we expect the estimated hazard ratio of *After-Split Dummy* to be greater than one in all specifications. We find that the estimated hazard ratio for *After-Split Dummy* is indeed greater than one and statistically significant at the 1% level for all specifications (see Panel A of Table 5). The estimated hazard ratio for *After-Split Dummy* in model (1) is 1.161, indicating that investors are 16.1% more likely to sell a stock in the six months after its split, controlling for other stock characteristics. Our results are robust across different event windows: the hazard ratio for *After-Split Dummy* takes on a statistically significant value of 1.171 for the 9-month window analysis in column (4), and 1.175 for the 12-month window analysis in column (7). These results suggest that investors are 16.1% to 17.5% more likely to sell their stock holding within the first year after the split.

Next, we repeat our analyses conducted in Panel A of Table 5 for reverse splits. Reverse splits are much rarer compared to forward splits. In these analyses, we exclude forward splits. Panel B of Table 5 reports the estimated results of the dynamic hazard regression using reverse split events instead of forward split events. All regression models in the table control for calendar year and month specific effects as well as for stock characteristics (i.e., size, book-to-market, and momentum). Models (2), (3), (5), (6), (8), and (9) further control for post reverse split returns, while models (3), (6), and (9) control for stock price values at the time of sale to account for potential clientele effects.

Given the increase in transaction costs after reverse stock splits, we would expect households to increase their holding periods, and thus we would expect the estimated hazard ratio on *After-R-Split dummy* to be less than one. In all specifications, we find that the estimated hazard ratio for *After-R-Split dummy* is less than one and economically and statistically highly significant. For example, the estimated hazard ratio for *After-R-Split dummy* in model (1) is 0.491, indicating that investors are 50.9% less likely to sell a stock in the six months after its reverse split. In Panel B, we exclude forward splits from our analyses, as including them would artificially strengthen our findings.

5.2 AMEX Tick Size Change

On September 3, 1992, the American Stock Exchange (AMEX) reduced its minimum price increment from 1/8th of a dollar to 1/16th of a dollar for stocks priced between \$1 and \$5. One of the motivations for this change was to reduce bid-ask spreads. A number of papers document that both quoted and effective spreads declined subsequent to this change (Ahn, Cao, and Choe 1996; Crack 1996). This quasi-exogenous shock to transaction costs presents us with another opportunity to assess our main hypothesis and address potential endogeneity issues.

To investigate the impact of this event, we use a similar approach as we used above to examine the impact of stock splits on investors' holding periods. We examine event windows of 6 months, 9 months, and 12 months after the implementation of the new tick size rule. We create a dummy variable (*After-AMEX tick change dummy*) that takes on a value of one if a month falls within the 6-month, 9-month or 12-month period.

Unlike splits, which are staggered over time, the AMEX tick size change event occurred at a single point in time. To control for potential confounding market-wide factors, we compare the change in the affected stocks that were priced \$1 to \$5 listed on AMEX to three groups of control stocks. That is, in addition to calculating the change in the holding period for stocks affected by the tick size change (i.e., treated) before and after the event, we also calculate the change in holding period for non-affected stocks (i.e., control) during the same time period. We then compare the change in the holding period for treated firms to the change in the holding period for non-treated (control) firms. The first control group contains firms that were priced between $1 and $5 but listed on the NYSE and NASDAQ exchanges and as such were not affected by the tick size change. The second group contains all firms on AMEX that were priced $5 or more. The third group contains all non-affected stocks on the three major exchanges.

The dynamic hazard regression results controlling for stock characteristics are reported in Table 6. Panels A, B, and C report the results for 6-, 9- and 12-month event windows, respectively. The coefficients reported under the row heading "*Treated*" are the hazard ratios estimated from the hazard regression on the interaction of the Treated group indicator with the *After-AMEX tick change dummy*, *Treated * After-AMEX tick change dummy*, while the coefficients reported under the row heading "Control" refer to the coefficient from the hazard regression on the interaction of the Control group indicator with the *After-AMEX tick change dummy*, *Control * After-AMEX tick change dummy*. The row with the header "Treated - Control" reports the difference in hazard rates between these two sets of interaction variables. In columns (1) of each of Panels A, B, and C, we focus only on firms listed on AMEX that were priced between $1 and $5 and thus directly impacted by the tick-size rule change. Specifically, we investigate the holding period decisions of investors for these impacted stocks and find that investors' likelihood of selling their impacted shares significantly increased after the tick-size reduction. For example, in Panel A, we observe that the estimated hazard ratio for the *After-AMEX tick change dummy* in model (1) is 1.167, indicating that investors were 16.7% more likely to sell an impacted stock (AMEX stock priced $1 to $5) in the 6-month period after the tick-size change. The results are qualitatively similar in Panels B and C, using 9-month (1.164) and 12-month (1.125) event windows, respectively.

In columns (2) through (4) for Panels A, B, and C, we investigate the differential impact of the tick-size rule change on investors' holding periods of impacted stocks (i.e., AMEX stocks with prices between \$1 and \$5) as well as those of different sets of control firms. The regression includes controls for size, book-to-market, momentum, and unrealized returns. In column (2), we use stocks that are similarly priced (with prices between \$1 to \$5) but listed on the NYSE and NASDAQ exchanges that were not affected by the tick size change as the control group. We find that sale probabilities of investors in treated stocks increase significantly by 18.7%, 20.5%, and 17.4%, respectively, in Panels A, B, and C after the tick size change. However, the sales probabilities of stocks priced similarly but listed on NYSE and NASDAQ increase by much smaller magnitudes, specifically, 6.7%, 12.6%, and 14.1%, respectively, in the 6, 9, and 12 months after the event. The differences between the coefficients in columns (2) of Panels A, B, and C for treated and control groups are always statistically significant at the 1% level, with meaningful economic differences.

In column (3), we compare the differential impact on holding periods for treated firms versus all other AMEX unaffected stocks. Similarly, we find that sale probabilities of retail investors in treated stocks increase by 12.5% throughout Panels A, B, and C, while sale probabilities of the rest of the AMEX stocks only increase by 2.6%. 3.7%, and 3.4% during the 6-, 9-, and 12-month windows, respectively. The differences in increases of sales probabilities are again significant both economically as well as statistically across the three event windows.

Finally, in column (4), we investigate the differential impact of the tick-size change on retail investors' holding periods for treated firms versus all other non-treated stocks (including all stocks on NYSE and NASDAQ as well as AMEX stocks that are priced more than \$5). We find that sale probabilities of retail investors in treated stocks increase by 11.8%, 12.9%, and 9.4%, respectively, while those for control stocks only increase by 1%, 7.3%, and 4.3%, respectively, for the 6-, 9-, and 12-month windows. The differences in increases of sales probabilities between treated and control stocks again are significant both economically and statistically across all three windows.

Overall, we find that the tick-size reduction leads retail investors to reduce their holding periods of treated firms in reaction to reduced transaction costs, consistent with our main hypothesis.

5.3 Finland Transactions

There may be sample selection concerns as the US sample covers only a subset of individual investors. To address this concern, we replicate our analyses using a transaction-level dataset from Finland that covers complete trading records of all individual investors between 1995 and 2003.

Table 7 reports the summary statistics for the Finnish stock and investor characteristics. Summary statistics are calculated by pooling annual observations over the 1995-2003 time period. All liquidity measures are calculated as described in section 3.2. The results show that our main transaction cost measure – adjusted Amihud ratio (*AdjIlliq*) – is positively skewed, with a mean of 10.61 and a smaller median of 6.21. Other transaction cost measures show a similar pattern. For example, *Zerofreq* has a mean of 21.90% and a median of 20.64%. Finally, we estimate *Closing price Spread (%)* following Barber and Odean (2000). The mean *Closing price Spread (%)* is 0.083, while the median is close to 0.[17] Size is also positively skewed, with the average market capitalization approximately 10 times as large as the median one.

The mean (median) investor age is 39.5 (40). About 33% of the primary account holders are female. The mean (median) household stock portfolio value is 10,823 (2,079) Euros in Finland. The mean (median) portfolio concentration is 0.20 (0.17), roughly corresponding to holding five stocks with equal value weights of 20%. Furthermore, 4% of households have traded options at least once and less than 1% have ever held a short position during the 1995-2003 time period. This is not surprising since Bris, Goetzmann, and Zhu (2007) suggest that short selling became legal in Finland in 1998 but that tax laws inhibit would-be short sellers.

[17] Brokerage commissions are not available for Finland.

We use a similar framework to the one we utilize for the US to test the validity of hypotheses (1) and (2) for Finnish investors. We run the same hazard regression, modeling the conditional probability that a stock is sold controlling for stocks' transaction costs, firm characteristics, available investors' demographic information and trade-related characteristics. Consistent with standard reporting convention, we report estimated hazard ratios from the hazard regressions instead of estimated hazard coefficients in Table 8.

The results estimated from the transaction-level Finnish dataset in Panel A of Table 8 are remarkably similar to our findings for the US. In the baseline model in column (1), the hazard ratio of the adjusted Amihud illiquidity measure (*AdjIlliq*) is 0.984, less than one and statistically significant. This indicates that if transaction costs (*AdjIlliq*) increase by one standard deviation (10.25), the investor is 15.2% less likely to sell that stock. We obtain comparable results using *Zerofreq* in column (2) and the *Closing price Spread (%)* in column (3). After we control for household- and firm-specific effects using stratification in column (4), the estimated hazard ratio of the adjusted Amihud illiquidity measure (*AdjIlliq*) is still less than one (0.976) and statistically significant. This coefficient indicates that with one standard deviation increase in *AdjIlliq*, the representative investor is 22% less likely to sell.

To explore how stock, investor and trade characteristics affect holding periods, we include additional controls in the regressions reported in columns (5) and (6) of Table 8. Controlling for stock characteristics (i.e., *Size*, *B/M*, *Momentum* and *Unrealized Returns*) in addition to household-specific effects in column (5) yields a statistically significant hazard ratio of less than one (0.979) for the adjusted Amihud illiquidity measure (*AdjIlliq*). We further control for both investor and stock characteristics available in the dataset in column (6). The estimated hazard ratio for the adjusted Amihud illiquidity measure (*AdjIlliq*) is statistically significant at 0.988, suggesting that the average investor is 11.6% less likely to sell a stock when the stock's transaction cost increases by one standard deviation.

The hazard ratios for investor characteristics are also quite similar to those for the US sample. Specifically, the hazard ratio for age is less than one, implying that older investors have lower turnover. In contrast, the hazard ratio for the male dummy is larger than one, suggesting that male investors tend to have shorter holding periods and trade more frequently. The hazard ratios for all trade-related variables are larger than one, suggesting that investors who trade options, who invest more capital in the stock market, and who concentrate their investments in fewer securities have shorter holding periods, consistent with our findings in the US data.

The loadings on stock characteristics are also similar to those in the US except for size. Similar to US investors, Finnish investors are also more likely to sell past winners while holding value stocks for longer periods. Unlike in the US, investors in Finland do not prefer to hold larger firms for longer periods. Altogether, the results in Panel A of Table 8 are similar to our US findings reported in Table 2. Individual investors in Finland are also cognizant of and pay attention to transaction costs when they make trading decisions.

In Panel B of Table 8, we further investigate heterogeneity in the relationship between transaction costs and holding periods. In particular, we examine if financially more sophisticated investors pay more attention to transaction costs. As in the US analysis, we assume that financial sophistication is correlated with education and resources available to each investor. We construct a similar sophistication measure. *Sophistication score* increases by one for each of the following three criteria met: if the investor is ranked among the top 25% based on the amount of capital invested in the stock market; if the investor has experience trading options (i.e., *Option User*

Dummy = 1), or if the investor's portfolio concentration is above that of the median investor.[18] Since the Finland transaction data do not provide information regarding investors' income, their professions, or whether the investor has ever traded any foreign securities, we exclude these criteria in the construction for the Finnish sophistication measure. Our sophistication measure for Finland ranges from a minimum of zero for the least sophisticated investors to a maximum of three for the most sophisticated investors.

We then divide Finnish investors into two subgroups based on their financial sophistication. Group 1 is comprised of the less sophisticated Finnish investors with *Sophistication Score* values of 0 or 1, while Group 2 includes the more sophisticated investors in Finland with *Sophistication Score* values of 2 or 3. Column (1) of Panel B in Table 8 reports that the hazard ratio of the Amihud illiquidity measure (*AdjIlliq*) for Group 1 is 0.992 and column (2) reports that the hazard ratio of *AdjIlliq* for Group 2 investors is 0.987. Both hazard ratios are statistically significant and less than one. The hazard ratio for the more sophisticated investors is smaller in magnitude compared to the hazard ratio for the less sophisticated investors.

These results suggest that the Finnish investors who are more sophisticated hold stocks with higher transaction costs for a longer period of time than less sophisticated investors, consistent with more financially sophisticated investors paying more attention to transaction costs compared to their less sophisticated peers. In both columns we control for *Size*, *B/M*, *Momentum*, *Unrealized Returns*, *Age*, and *Male Dummy*. Using alternative transaction cost measures, including the proportion of zero return days, (*Zerofreq*), as well as actual transaction costs (*Closing Price Spread (%)*), generates comparable results. Overall, these findings suggest that our findings in the US are unlikely to be driven by the specific sample of investors and the time period we study.

[18] Since a small percentage of Finnish households have ever held short positions, we do not include this variable in the construction of our sophistication measure.

6. Conclusion

This paper investigates how the trading decisions of 66,000 households in the US and 303,000 households in Finland are influenced by transaction costs. Two main conclusions follow from our analyses. First, we show that transaction costs are an important determinant of the investment decisions of individual investors. Consistent with theoretical models of investor behavior, households rationally reduce the frequency with which they trade illiquid securities that are subject to high transaction costs. This finding is robust to controlling for household and stock characteristics. Second, we show that there is cross-sectional variation in the relationship between holding periods and transaction costs across households. Particularly, the relationship between transaction costs and holding periods is stronger among more sophisticated investors.

To address endogeneity and selection concerns, we examine how investors' holding periods change around quasi-exogenous changes in transaction costs. We find that investors shorten their holding periods after stock split events in response to stock liquidity increases. We also document similar declines in holding periods after an exogenous reduction in tick size for stocks priced under $5 at AMEX in 1992.

Overall, our findings challenge the notion that individual investors ignore non-salient costs when making investment decisions. We show that individual investors are cognizant of at least one type of non-salient cost, namely the cost of trading stocks, revealing a unique aspect of their rationality.

Recognizing the dynamic nature of financial markets, our study acknowledges the significant shifts that have occurred since the conclusion of the data used in our analysis. In the last two decades, there has been a notable increase in households' access to financial information and a substantial reduction in transaction costs due to technological developments. Despite these changes, there is still significant variation in aggregate liquidity over time as well as cross-sectional variation in liquidity across stocks. Events such as the global financial crisis and the onset of the COVID-19 pandemic have highlighted how liquidity can rapidly diminish during crises. Furthermore, the COVID-19 pandemic era witnessed a surge in active trading by households, particularly in so called 'meme' stocks, reigniting discussions about investor biases, rational decision-making, and liquidity risks. These developments underscore the relevance and the importance of results presented in this study and also the need for future research to explore how the emerging generation of retail investors factors transaction costs into their investment decisions.

References

Acharya, V. V., and L. H. Pedersen, 2005, “Asset pricing with liquidity risk,” Journal of Financial Economics, 77, 375-410

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

Alexander, G., J. Jones, and P. J. Nigro, 1998, “Mutual fund shareholders: Characteristics, investor knowledge, and sources of information,” Financial Services Review, 7, 301-316

Allison, P. D., and N. A. Christakis, 2006, “Fixed-effects methods for the analysis of nonrepeated events,” Sociological Methodology, 36, 155-172

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

Amihud, Y., and H. Mendelson, 1986, “Asset pricing and the bid-ask spread,” Journal of Financial Economics, 17, 223-249

Anginer, D., S. X. Han, and Ç. Yıldızhan, 2024 “Online Appendix to Accompany “Do Individual Investors Ignore Transaction Costs?” available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4700788, SSRN

Applied Research and Consulting LLC, 2003, “NASD Investor Literacy Research,” available at: http://www.finrafoundation.org/surveyexecsum.pdf

Atkins, A. B., and E. A. Dyl, 1997, “Transactions costs and holding periods for common stocks,” Journal of Finance, 52(1), 309-325

Bali, T. G., N. Cakici, and R. F. Whitelaw, 2011, “Maxing out: Stocks as lotteries and the cross-section of expected returns,” Journal of Financial Economics, 99(2), 427-446

Bali, T. G., R. F. Engle, and Y. Tang, 2017, “Dynamic conditional beta is alive and well in the cross section of daily stock returns,” Management Science, 63(11), 3531-3997

Barber, B. M., Y. Lee, Y. Liu, and T. Odean, 2009, “Just how much do individual investors lose by trading?” Review of Financial Studies, 22(2), 609-632

Barber, B. M., and T. Odean, 2000, “Trading is hazardous to your wealth: The common stock investment performance of individual investors,” The Journal of Finance, 55, 773-806

Barber, B. M., and T. Odean, 2001, “Boys will be boys: Gender, overconfidence, and common stock investment,” The Quarterly Journal of Economics, 116(1), 261-292

Barber, B. M., T. Odean, and L. Zheng, 2005, “Out of sight, out of mind: The effects of expenses on mutual fund flows,” Journal of Business, 78, 2095-2119

Barber, B. M., T. Odean, and N. Zhu, 2006, “Do noise traders move markets?” working paper, University of California, Davis

Bris, A., W. Goetzmann, and N. Zhu, 2007, “Efficiency and the bear: Short sales and markets around the world,” The Journal of Finance, 62, 1029-1079

Byun, J., and M. S. Rozeff, 2003, “Long-run performance after stock splits: 1927 to 1996,” The Journal of Finance, 58, 1063-1085

Chetty, R., A. Looney, and K. Kroft, 2009, “Salience and taxation: Theory and evidence,” American Economic Review, 99(4), 1145-1177.

Christakis, N.A. and P. D. Allison, 2006, “Mortality after the hospitalization of a spouse,” New England Journal of Medicine, 354, 719-730

Conroy, R. M., R. S. Harris, and B. A. Benet, 1990, "The effects of stock splits on bid-ask spreads," The Journal of Finance, 45, 1285-1295

Constantinides, G. M., 1986, "Capital market equilibrium with transaction costs," Journal of Political Economy, 94, 842-862

Coval, J. D., D. A. Hirshleifer, and T. Shumway, 2005, "Can individual investors beat the market?" working paper, University of Michigan

Cox, D., 1972, "Regression models and life-tables," Journal of the Royal Statistical Society, Series B (Methodological), 34(2), 187-220

Cox, D., and D. Oakes, 1984, Analysis of survival data, Chapman and Hall, London; New York

Crack, T., 1996, "Tinkering with ticks: Choosing minimum price variation for US equity markets," working paper, University of Otago

Datar, V. T., N. Y. Naik, and R. Radcliffe, 1998, "Liquidity and stock returns: An alternative test," Journal of Financial Markets, 1(2), 203-219

Desai, A. S., M. Nimalendran, and S. Venkataraman, 1998, "Changes in trading activity following stock splits and their effect on volatility and the adverse-information component of the bid-ask spread," Journal of Financial Research, 21,159-183

Feng, L., and M. S. Seasholes, 2005, "Do investor sophistication and trading experience eliminate Behavioral biases in financial markets?" Review of Finance, 9, 305-351

Finkelstein, A., 2009, "E-ztax: Tax salience and tax rates," Quarterly Journal of Economics, 124(3), 969-1010

French, Kenneth R., 2008, "Presidential address: The cost of active investing," The Journal of Finance, 63(4), 1537-1573

Gil-Bazo, J., and P. Ruiz-Verdu, 2008, "When cheaper is better: Fee determination in the market for equity mutual funds," Journal of Economic Behavior & Organization, 67(3), 871-885

Gil-Bazo, J., and P. Ruiz-Verdu, 2009, "The relation between price and performance in the mutual fund industry," The Journal of Finance, 64(5), 2153-2183

Goetzmann, W. N., A. Kumar, 2008, "Equity portfolio diversification," Review of Finance 12(3), 433-463

Grinblatt, M., and M. Keloharju, 2000, "The investment behavior and performance of various investor types: a study of Finland's unique dataset," Journal of Financial Economics, 55(1), 43-67

Grinblatt, M., and M. Keloharju, 2001, "What makes investors trade?" The Journal of Finance, 56, 589-616

Hasbrouck, J., 2004, "Liquidity in the future pits: Inferring market dynamics from incomplete data," Journal of Financial and Quantitative Analysis, 39(2), 305-326

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

Heaton, J., and D. J. Lucas, 1996, "Evaluating the effects of incomplete markets on risk sharing and asset pricing," Journal of Political Economy, 104(3), 433-487

Hossain, T., and J. Morgan. 2006, "… Plus shipping and handling: Revenue (non) equivalence in field experiments on eBay," Advances in Economic Analysis and Policy, 6(2), 1-27.

Hu, S., 1997, "Trading turnover and expected stock returns: The trading frequency hypothesis and evidence from Tokyo Stock Exchange," working paper, National Taiwan University

Huang, M., 2003, “Liquidity shocks and equilibrium liquidity premia,” Journal of Economic Theory, 109, 104-129

Hvidkjaer, S., 2006, “A trade-based analysis of momentum,” The Review of Financial Studies 19(2), 457-491

Ivkovic, Z., J. Poterba, and S. J. Weisbenner, 2005, “Tax-motivated trading by individual investors,” American Economic Review, 95(5), 1605-1630

Ivkovic, Z., C. Sialm, and S. J. Weisbenner, 2008, “Portfolio concentration and the performance of individual investors,” Journal of Financial and Quantitative Analysis, 43(3), 613-655

Kryzanowski, L., and H. Zhang, 1996, “Trading patterns of small and large traders around stock split ex-dates,” The Journal of Financial Research, 19(1), 75-90

Kumar, A., 2009, “Who gambles in the stock market?” The Journal of Finance, 64(4), 1889-1933

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka, 1999, “A new estimate of transaction costs,” Review of Financial Studies, 12(5), 1113-1141

Lin, D. Y. and L. J. Wei, 1989, “The robust inference for the Cox proportional hazards model,” Journal of the American Statistical Association, 84, 1074-1078

Lo, A. W., H. Mamaysky, and J. Wang, 2004, “Asset prices and trading volume under fixed transaction costs,” Journal of Political Economy, 112, 1054-1090

Lynch, A. W., and S. Tan, 2011, “Explaining the magnitude of liquidity premia: The roles of return predictability, wealth shocks and state dependent transaction costs,” Journal of Finance 66(4), 1329-1368

Odean, T., 1999, “Do investors trade too much?” American Economic Review, 89, 1279-1298.

Roll, R., 1984, “A simple implicit measure of the effective bid-ask spread in an efficient market,” Journal of Finance, 39, 1127-1139

Schultz, Paul, 2000, “Stock splits, tick size, and sponsorship,” Journal of Finance, 55, 429-450

Seru, A., T. Shumway, and N. Stoffman, 2010, “Learning by trading”, Review of Financial Studies 23(2), 705-73

Vayanos, D., 1998, “Transactions costs and asset prices: A dynamic equilibrium model,” Review of Financial Studies, 11, 1-58

Vayanos, D., and J.C. Vila, 1999, “Equilibrium interest rate and liquidity premium with transaction costs,” Economic Theory, 13, 509-539

Figure 1: Survival Probabilities for Stocks in the United States and Finland

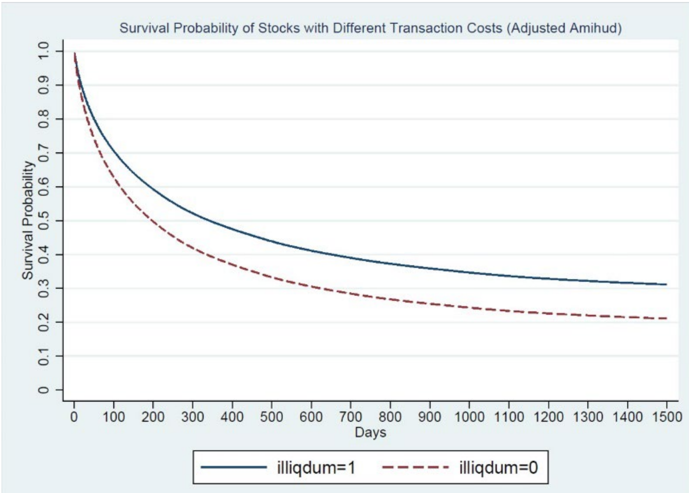

Figure 1a plots the Kaplan-Meier survival probabilities for two groups of stocks held by households in the United States over the 1991-1996 time period. Illiquid stocks in the figure are stocks that belong to the top decile based on their adjusted Amihud illiquidity measure. The solid line represents the probability of holding onto these illiquid stocks, and the dashed line represents the probability of holding all of the other stocks.

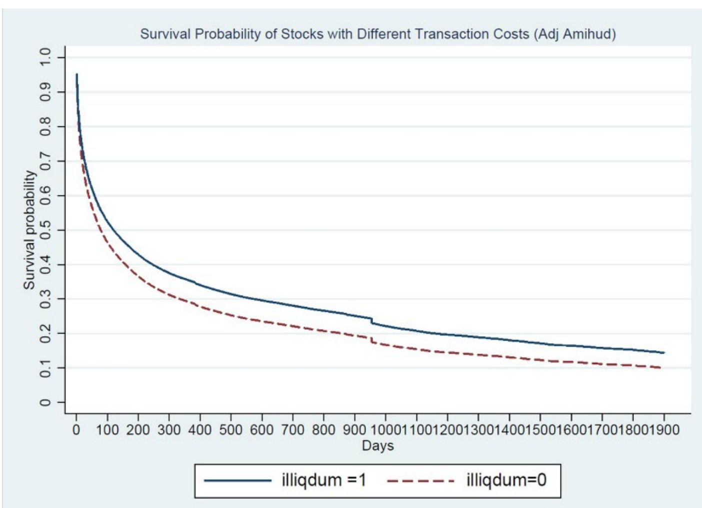

Figure 1b plots the Kaplan-Meier survival probabilities for two groups of stocks held by households in Finland over the 1995-2003 time period. Illiquid stocks in the figure are stocks that belong to the top decile based on their adjusted Amihud illiquidity measure. The solid line represents the probability of holding onto these illiquid stocks, and the dashed line represents the probability of holding all of the other stocks.

Table 1: Summary Statistics of Stock and Investor Characteristics in the US

Table 1 reports the descriptive statistics for stock and investor characteristics in the US. Summary statistics are calculated by pooling annual observations over the 1991-1996 time period. Price is the annual average of daily closing prices. Market Cap is the average market capitalization in millions of US dollars. B/M is the book-to-market ratio. Past Returns (-12, -2) is a proxy for momentum and measures cumulative returns during the past 10 months starting at month -12 and ending two months prior. AdjIlliq is the adjusted Amihud illiquidity ratio. Zerofreq is the proportion of zero-return days which is calculated as the percentage of zero-return days within a year. All liquidity measures are annual averages as defined in the text. Closing price spread is calculated as the closing price divided by the purchase price on the day of the transaction minus one, and then multiplied by minus one. Commission is calculated as the commission paid divided by the value of the purchase. Effective spread and relative spread are calculated using TAQ data. Effective spread / Price is defined as transaction price minus bid-ask midpoint then multiplied by two and scaled by price. Relative spread / Price is defined as NBBO quoted spread divided by price, and Relative spread / Mid is defined as quoted spread scaled by the bid-ask midpoint. Depth is defined as the midpoint of bid size and offer size (both in number of round lots). Panel A reports the characteristics only for stocks that have observed individual investor transactions in the dataset, while Panel B reports the stock characteristics of the CRSP universe during the same period. As closing price spread and commission are calculated using transaction level data, we can only compute them for stock included in our transaction database, not for the entire CRSP stock universe. Differences in means between the CRSP sample and the sample used in the analyses are reported in the last column of Panel B. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively. Panel C reports the characteristics of investors included in the dataset. Age in 1996 is the biological age of the investor in 1996. Married Dummy is a dummy variable that equals one for married traders. Male Dummy is a dummy variable that equals one if the head of household is male. Professional Dummy is a dummy variable equal to one for traders that hold either technical or managerial positions. Retired Dummy is a dummy variable that is equal to one for traders who already retired. Retirement Acct Dummy is a dummy variable that equals one if the transaction takes place in a retirement account such as a 401(k). Portfolio Concentration is calculated following Ivkovic, Sialm, and Weisbenner (2008) as the sum of squared value weights of each stock in a household's portfolio. Equity Portfolio Value reports the total dollar value of an investor's equity portfolio every month. Income is annual self-reported income in thousands of dollars. Option User Dummy is a dummy variable that equals one if a trader has traded options at least once over the entire sample period. Foreign Securities Dummy is a dummy variable that equals one if a trader has traded any foreign assets, including ADRs, foreign stocks, or foreign mutual funds, at least once over the entire sample period. Short User Dummy is a dummy variable that equals one if an investor has shorted any security at least once over the entire sample period.

	Mean	P25	Median	P75	Std. Dev
Panel A:	Sample Stock Characteristics				
Price ($)	20.51	4.75	11.88	23.50	308.55
Market Cap ($M)	896.11	25.14	87.91	364.45	4182.50
B/M	0.78	0.30	0.57	0.94	2.94
Past Returns (-12, -2)	0.25	-0.10	0.08	0.30	2.11
AdjIlliq	5.04	0.36	1.18	6.26	7.49
ZeroFreq (%)	7.14	0.00	4.86	10.42	8.80
Closing price Spread % + Commission (%)	2.12	0.53	1.69	3.37	2.95
Effective Spread/ Price (%)	1.84	0.52	0.90	1.72	3.25
Relative Spread/ Price (%)	2.95	0.98	1.56	2.83	4.69
Relative Spread/Mid (%)	2.92	0.98	1.56	2.82	4.48
Log (1+depth)	3.25	2.61	3.20	3.87	0.94

	Mean	P25	Median	P75	Std. Dev	Difference
Panel B:	CRSP Stock Characteristics					
Price ($)	20.19	4.75	11.88	23.25	298.71	0.33
Market Cap ($M)	850.68	23.20	80.63	336.22	4057.39	45.43
B/M	0.72	0.30	0.56	0.91	0.89	0.06***
Past Returns (-12, -2)	0.24	-0.10	0.08	0.29	2.12	0.004
AdjIlliq	5.56	0.38	1.36	7.55	7.92	-0.52***
ZeroFreq (%)	7.53	0.00	4.86	11.11	9.25	-0.39***
Effective Spread/ Price (%)	2.51	0.66	1.10	2.12	4.37	-0.67***
Relative Spread/ Price (%)	3.91	1.12	1.74	3.34	6.45	-0.96***
Relative Spread/Mid (%)	3.85	1.12	1.74	3.33	6.22	-0.93***
Log (1+depth)	3.69	3.05	3.60	4.24	1.07	-0.44***

Panel C:	Investor Characteristics				
Age in 1996	49.58	40	48	58	12.40
Married Dummy (1 = married)	0.76	1	1	1	0.43
Male Dummy (1 = male)	0.90	1	1	1	0.30
Professional Dummy	0.66	0	1	1	0.47
Retired Dummy	0.15	0	0	0	0.36
Retirement Acct Dummy	0.39	0	0	1	0.49
Portfolio Concentration	0.52	0.28	0.48	0.73	0.28
Equity Portfolio Value ($)	80,342	8,900	22,952	62,087	313,568
Income ($K)	76.84	45	87.5	112.5	33.19
Option User Dummy	0.14	0	0	0	0.34
Foreign Securities Dummy	0.22	0	0	0	0.42
Short User Dummy	0.38	0	0	1	0.49

Table 2: Impact of Transaction Costs on Households' Holding Periods in the US, Hazard Analysis

This table examines the impact of transaction costs on individual investors' holding periods in the US between 1991 and 1996 using a hazard model framework. The conditional probability of sale is the dependent variable. Independent variables include the adjusted Amihud illiquidity ratio, firm characteristics, a set of demographic controls, and a variety of trade variables. Proxies for transactions costs (AdjIlliq) are averaged over the previous 12 months before each sale transaction. Size is measured as the log of the market capitalization at the end of the month prior to the sale transaction. B/M or book-to-market ratio is computed as the ratio of previous year-end book value to the most recent market capitalization. Momentum is the cumulative returns over the ten-month period from month -12 to month -2. Stock characteristics are calculated at the beginning of the month when a sale takes place. Unrealized returns are calculated using the price differentials observed at the time of closing of the position and the time of purchase, divided by initial investment made at the time of purchase. For positions not closed at the end of the sample period, we assume the price at the last day of our sample period as the closing price. Beta is CAPM beta estimated each month following Bali, Cakici, and Whitelaw (2011). Ivol is idiosyncratic volatility, estimated as the monthly standard deviation of a stock's daily return residual from the Fama-French three-factor model. MarPrc is the maximum daily return during the month prior to sales. Age refers to the age of the head of household. Income is the total self-reported annual income. Married Dummy is a dummy variable that equals one if the investor is married. Male Dummy is equal to one if the head of household is a male. Professional Dummy is one for investors who hold technical or managerial positions, and Retired Dummy is equal to one for investors who already retired. Retirement Acct Dummy equals one if the transaction account is a retirement (IRA or Keogh) account. Trade variables for each individual investor are derived from all the transactions he/she executes during the sample period. Short User Dummy equals one if an investor executed at least one short sale during the sample period. Option User Dummy is one if an investor ever traded options. Foreign Securities Dummy is set to one if an investor traded at least once any foreign assets, including ADRs, foreign stocks, or foreign mutual funds during the sample period. Log (Equity Portfolio Value) is the logarithmic value of the household's total equity holdings every month. Portfolio Concentration is defined as in Ivkovic, Sialm, and Weisbenner (2008) and is equal to the sum of squared value weights of each stock in a household's portfolio. Calendar month dummies (not reported) are twelve dummy variables that equal one if the sale transaction happens during the specific month. Year dummies (not reported) equal one for the year during which a transaction happens. Clustered robust standard errors are calculated at the household level. Robust standard errors are adjusted as in Lin and Wei (1989). Ties are handled using the Efron procedure. The Wald test is used for each additional set of regressors. The *p*-values are reported below each coefficient. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
AdjIlliq	0.981***	0.973***	0.981***	0.975***	0.975***	0.977***
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
			Stock characteristics			
Size			0.649***	0.951***	0.950***	0.962***
			<.0001	<.0001	<.0001	<.0001
B/M			0.681***	0.986***	0.988***	1.008*
			<.0001	<.0001	<.0001	0.078
Momentum			1.135***	1.097***	1.094***	1.093***
			<.0001	<.0001	<.0001	<.0001
Unrealized Returns				1.134***	1.138***	1.140***
				<.0001	<.0001	<.0001
Beta						1.111***
						<.0001
Ivol						2.807***
						<.0001
MaxPrc						0.999
						0.477
			Demographic variables			
Age				0.996***	0.997***	0.996***
				<.0001	<.0001	<.0001
Log (Income)				0.999***	0.999***	0.930***
				<.0001	<.0001	<.0001
Married Dummy				0.955***	0.960***	0.963***
				<.0001	<.0001	<.0001
Male Dummy				1.113***	1.110***	1.115***
				<.0001	<.0001	<.0001
Professional Dummy				1.020**	1.020*	1.019*
				0.050	0.051	0.063
Retirement Acct Dummy				0.885***	0.893***	0.906***
				<.0001	<.0001	<.0001
Retired Dummy				1.077***	1.076***	1.075***
				<.0001	<.0001	<.0001
			Trade Characteristics			
Foreign Securities Dummy				1.233***	1.237***	1.224***
				<.0001	<.0001	<.0001
Option User Dummy				1.484***	1.464***	1.430***
				<.0001	<.0001	<.0001
Short User Dummy				1.983***	1.980***	1.909***
				<.0001	<.0001	<.0001
Log (Equity Portfolio Value)				1.001***	1.000***	1.083***
				<.0001	<.0001	<.0001
Portfolio Concentration				2.589***	2.527***	3.107***
				<.0001	<.0001	<.0001
Firm stratification	No	Yes	No	No	No	No
Household stratification	No	Yes	No	No	No	No
Calendar month/year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Remove day trading	No	No	No	No	Yes	No
Observations	799,469	799,469	589,794	115,147	113,593	115,147
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Table 3: Alternative Transaction Costs on Households' Holding Periods in the US, Hazard Analysis

This table examines the impact of transaction costs on individual investors' holding periods in the US between 1991 and 1996 using a hazard model framework. The conditional probability of sale is the dependent variable. Independent variables include six alternative measures of illiquidity. In Panel B, the independent variables also include a set of firm characteristics, demographic controls, and trade variables. Proxies for Transaction costs are averaged over the previous 12 months before each sale transaction. All transaction cost measures are defined as in Table 1. As depth tends to be skewed, we use log(1+depth) in our analysis. All of the control variables are as defined in Table 2. Calendar month dummies (not reported) and Year dummies (not reported) are included in all specifications. Robust standard errors are adjusted as in Lin and Wei (1989). Ties are handled using the Efron procedure. The Wald test is used for each additional set of regressors. The *p*-values are reported below each coefficient. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Illiquidity measure	ZeroFreq	Closing price Spread % + Commission (%)	Effective Spread/Price (%)	Relative Spread/Price (%)	Relative Spread/Mid (%)	Log (1+depth)
Illiquidity	0.322***	0.945***	0.988***	0.993***	0.993***	0.916***
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Firm stratification	No	No	No	No	No	No
Household stratification	No	No	No	No	No	No
Calendar month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Calendar year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	766,168	778,052	616,825	616,825	616,825	536,772
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Table 4: Transaction Costs and Holding Periods for Investors of Various Sophistication

This table examines how the impact of transaction costs on individual investors' holding periods in the US differ across investors with various sophistication levels. Investor sophistication is presumed to cumulatively increase with each of the following seven criteria met: if the investor has an income higher than $75K; if the investor is ranked among the top 25% of all investors based on equity holdings at any point in time during the sample period; if the investor holds either technical or managerial positions and as such is considered a professional; if the investor traded options at least once during the entire sample period; if the investor has ever held any short positions during the sample period; if the investor has ever traded foreign securities, including ADRs, foreign stocks or mutual funds; and if the investor's portfolio is more concentrated than the median investor's (i.e., if the investor's portfolio concentration is greater than 0.48). The most sophisticated investors in the US have a sophistication score of 7, while the least sophisticated have a sophistication score of 0. We categorize all households into three subsamples based on their sophistication. Group 1 includes the least sophisticated investors, whose sophistication scores are between 0 and 2; Group 2 is for investors whose sophistication scores are between 3 and 5; and Group 3 contains the most sophisticated investors, whose sophistication scores are 6 or 7. We then estimate hazard regression where the conditional probability of sale is the dependent variable. Independent variables include adjusted Amihud illiquidity ratio, stock characteristics, and demographic controls. All variables are as defined in Table 2. Since we use Income, Professional Dummy, Short User Dummy, Option User Dummy, Foreign Securities Dummy, Log (Equity Portfolio Value), and Portfolio Concentration to calculate Sophistication, these variables are not included as independent variables in the analyses. Robust standard errors are adjusted as in Lin and Wei (1989). Ties are handled using the Efron procedure. The Wald test is used for each additional set of regressors. The *p*-values are reported below each coefficient. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Sophistication Group	(1)	(2)	(3)
Sophistication Score	0, 1, 2	3, 4, 5	6, 7
AdjIlliq	0.984***	0.975***	0.948**
	<.0001	<.0001	0.011
		Stock Characteristics	
Size	0.797***	0.677***	0.750***
	<.0001	<.0001	<.0001
B/M	0.963***	0.846***	0.793***
	<.0001	<.0001	<.0001
Momentum	1.083***	1.118***	1.102***
	<.0001	<.0001	<.0001
Unrealized Returns	1.131***	1.013	1.011
	<.0001	0.294	0.828
		Demographic Variables	
Age	0.998***	0.998***	0.981***
	0.002	<.0001	<.0001
Married Dummy	0.949***	0.883***	0.795***
	<.0001	<.0001	<.0001
Male Dummy	1.093***	1.116***	1.119
	<.0001	<.0001	0.181
Retirement Acct Dummy	0.870***	0.877***	0.777***
	<.0001	<.0001	<.0001
Retired Dummy	1.104***	1.294***	1.650***
	<.0001	<.0001	0.001
Firm stratification	Yes	Yes	Yes
Household stratification	No	No	No
Calendar month dummies	Yes	Yes	Yes
Calendar year dummies	Yes	Yes	Yes
Observations	81,685	76,894	8,215
Wald test	<.0001	<.0001	<.0001

Table 5: Impact of US Stock Splits on Holding Period Decisions

This table examines the impact of stock splits on individual investors' holding period decisions. It reports the estimated hazard ratios from dynamic hazard regressions in the US where the conditional probability of sale is the dependent variable. We employ three different event windows as defined before, specifically 6, 9, and 12 months after stock splits. For each stock-holding position, we need to have one observation for every day starting from the very first day the position is open, up to and including the day the stock is sold, or in cases where sales of stocks are not observed, until the last day of our sample period. After-Split Dummy (After-R-Split Dummy) is a dummy variable used for forward (reverse) splits that equals one if an observation falls in one of the three event windows. To efficiently estimate such a huge panel of data with likelihood function, we follow Allison and Christakis (2006) and split the duration period into multiple periods (i.e., pre-event, event, and post-event period). For each period there are multiple observations where the After-Split Dummy equals either 0 or 1; it is only the last observation with a distinct After-Split Dummy (After-R-Split Dummy) value for each period that matters in the estimation. Thus, we keep the last observation for each period with a distinct After-Split Dummy (After-R-Split Dummy) value. To address the concern that the results might be driven by post-split returns, we calculate returns for each period accordingly and control for these returns in models (2), (3), (5), (6), (8), and (9) in the table. Finally, we account for the possibility that stock splits may lead to clientele effects: forward (reverse) splits may attract clienteles that prefer lower-priced (higher-priced) equities. Columns (3), (6), and (9) address the clientele issue by controlling for the stock price at the time of sale. For positions not closed at the end of our sample period, we use the stock price on the last day of our sample period as the closing price. We report the estimated hazard ratios on the After-Split Dummy for forward splits in Panel A and the estimated hazard ratios on the After-R-Split Dummy for reverse-splits in Panel B. In all specifications, we control for size, book-to-market, and momentum as defined previously, as well as time-specific effects with calendar year and month dummies. For brevity, estimated hazard ratios for stock characteristics and calendar year and month dummies are not reported. The *p*-values are reported below each coefficient. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: Impact of Forward Splits on Holding Period Decisions

Window	6-Month			9-Month			12-Month		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After-Split Dummy	1.161***	1.087***	1.089***	1.171***	1.099***	1.099***	1.175***	1.100***	1.100***
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Observations	681,042	681,021	681,021	682,636	682,537	682,537	685,571	685,409	685,409
Stock controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Split Return Control	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Stock Price Control	No	No	Yes	No	No	Yes	No	No	Yes
Calendar Year & Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Panel B: Impact of Reverse Stock Splits on Holding Period Decisions

Window	6- Month			9-Month			12-Month		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
After-R-Split Dummy	0.491***	0.485***	0.480***	0.543***	0.544***	0.535***	0.628***	0.626***	0.536***
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Observations	386,490	386,230	374,001	387,461	386,836	374,314	387,197	386,843	374,353
Stock controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Split Return Control	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Stock Price Control	No	No	Yes	No	No	Yes	No	No	Yes
Calendar Year & Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Table 6: Holding Period Changes around September 3rd, 1992, AMEX Tick Size Changes

This table examines the impact of AMEX tick size changes on September 3rd, 1992, on individual investors' holding period decisions. It estimates changes in hazard ratios using dynamic hazard regressions. The treated group consists of AMEX stocks priced between $1 and $5 on the day the tick size change was implemented. We estimate the differential impact of this rule change on the holding period decisions in treated stocks versus stocks that were not impacted. We employ three alternative control groups in this analysis. Model (1) focuses only on stocks for which the tick size changed on September 3rd, 1992 (the treated group) and simply investigates the change in the likelihood of sale after the implementation of the tick size rule change compared to before the implementation of the rule change. Model (2) uses stocks priced between $1 and $5 and listed on NYSE or NASDAQ as the control group. Model (3) uses all other stocks listed on AMEX that are priced above $5 as the control group. Model (4) includes all stocks that are not in the treated group as the control group: this includes all stocks listed on NYSE and NASDAQ, as well as stocks in AMEX that are priced above $5. The conditional probability of sale is the dependent variable, and we employ three different event windows as defined before: 6, 9, and 12 months subsequent to September 3rd, 1992. We follow Allison and Christakis (2006) and separate the sample period into multiple sub-periods (i.e., pre-event, event, and post-event period). Specifically, the first period covers the time period from purchasing the share until the tick size change. In this period (pre-event), the After-AMEX tick change dummy equals zero. The second period is the time period from September 3rd, 1992, until the end of the event window of interest (i.e., 6, 9 and 12 months). In the second period, the After-AMEX tick change dummy equals 1. The third period corresponds to the time-period after the event (post-event window), for which the After-AMEX tick change dummy equals zero again. We estimate the hazard ratio of the After-AMEX tick change dummy for the treated group as well as for the control groups and then examine if the difference between the two estimated hazard ratios for the treatment group vs. the control group is significant. In all our analyses, we control for size, book-to-market, momentum, and unrealized returns. In all specifications, we use firm stratification to account for firm-specific factors. The table reports the estimated hazard ratios on the After-AMEX tick change dummy for the treated firms and untreated / control firms. Panels A, B, and C document the estimated results for the 6-, 9-, and 12-month event windows, respectively. For brevity, estimated hazard ratios for stock characteristics are not reported. The *p*-values are reported below each coefficient. We further report the Chi-square (χ^2) and *p*-value corresponding to testing the difference between the hazard ratios of the After-AMEX tick change dummy for the treated group vs. the control group. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: 6-Month Analysis				
	(1)	(2)	(3)	(4)
Control Group	None	NYSE & NASDAQ stocks priced [1,5]	AMEX Stocks Priced >$5	All stocks
	After-AMEX tick change dummy			
Treated	1.167^{**}	1.187^{***}	1.125^{***}	1.118^{***}
	0.037	<.0001	<.0001	<.0001
Control		1.067^{**}	1.026^{***}	1.010^{***}
		<.0001	<.0001	0.005
Treated – Control (χ^2)		1025.09	11.32	28.75
p-value		<.0001	0.0008	<.0001
Stock controls	Yes	Yes	Yes	Yes
Firm stratification	Yes	Yes	Yes	Yes
Observations	8,279	63,612	31,979	755,378
Wald test	<.0001	<.0001	<.0001	<.0001

Panel B: 9-Month Analysis				
	(1)	(2)	(3)	(4)
Control Group	None	NYSE & NASDAQ stocks priced [1,5]	AMEX Stocks Priced >$5	All stocks
	After-AMEX tick change dummy			
Treated	1.164**	1.205***	1.125***	1.129***
	0.014	<.0001	<.0001	<.0001
Control		1.126***	1.037**	1.073***
		<.0001	0.041	<.0001
Treated – Control (χ^2)		535.59	3.92	4.96
p-value		<.0001	0.0477	0.026
Stock controls	Yes	Yes	Yes	Yes
Firm stratification	Yes	Yes	Yes	Yes
Observations	8,343	64,017	32,216	761,808
Wald test	<.0001	<.0001	<.0001	<.0001

Panel C: 12-Month Analysis				
	(1)	(2)	(3)	(4)
Control Group	None	NYSE & NASDAQ stocks priced [1,5]	AMEX Stocks Priced >$5	All stocks
	After-AMEX tick change dummy			
Treated	1.125**	1.174***	1.125***	1.094***
	0.029	<.0001	<.0001	0.001
Control		1.141***	1.034	1.043***
		0.001	0.18	<.0001
Treated – Control (χ^2)		1108.91	3.16	4.41
p-value		<.0001	0.0704	0.0357
Stock controls	Yes	Yes	Yes	Yes
Firm stratification	Yes	Yes	Yes	Yes
Observations	8,398	64,346	32,303	767,109
Wald test	<.0001	<.0001	<.0001	<.0001

Table 7: Summary Statistics of Stock and Investor Characteristics in Finland

This table reports the descriptive statistics for stock and investor characteristics in Finland. Summary statistics are calculated by pooling annual observations over the 1995-2003 time-period. Price is the annual average of the daily closing prices. Market Cap is the average market capitalization in millions of Euros. AdjIlliq is the adjusted Amihud illiquidity ratio. Zerofreq is zero-return frequency, which reports the percentage of zero-return days. Following Barber and Odean (2000), spread is calculated as the purchase price divided by the closing price on the day of the transaction minus one, and then multiplied by minus one for purchase. Age in 1995 is the biological age of the investor in 1995. Male Dummy (1 = male) is a dummy variable that equals one for male traders. Portfolio concentration is defined as in Ivkovic, Sialm, and Weisbenner (2008) and is calculated as the sum of squared value weights of each stock in a household's portfolio. Equity Portfolio Value is the annual average market value of an investor's portfolio in Euros using daily closing prices. Option User Dummy is a dummy variable that equals one for traders that have traded options at least once over the entire sample period.

	Mean	P25	Median	P75	Std. Dev
	Stock Characteristics				
Price (€)	12.61	2.69	7.67	16.4	11.20
Market Cap (€M)	1132	33	125	498	8414.34
AdjIlliq	10.61	1.07	6.21	20.12	10.25
Zerofreq (%)	21.90	13.50	20.64	27.75	13.42
Spread (%)	0.083	-2.93	0	3.25	5.52
	Investor Characteristics				
Age in 1995	39.5	27	40	52	18.48
Male Dummy (1 = male)	0.67	0	1	1	0.47
Portfolio Concentration	0.20	0.09	0.17	0.27	0.18
Equity Portfolio Value (€)	10,823	1,341	2,079	5,292	80,125
Option User Dummy	0.04	0	0	0	0.18

Table 8: Impact of Liquidity on Households' Holding Periods in Finland, Hazard Analysis

This table examines the impact of stock liquidity on individual investors' holding periods in Finland using a hazard model framework similar to the one used in Table 2 for the US data. Panel A reports the estimated hazard ratios from hazard regressions where the conditional probability of sale is the dependent variable. Independent variables include transaction cost measures: the adjusted Amihud illiquidity ratio (alternatively Zerofreq or Closing Price Spread (%)), firm characteristics, and a set of demographic controls and trade variables. The variables are described in Table 2. Panel B investigates if sophistication affects an investor's attention to transaction costs. A Finnish investor's sophistication is presumed to cumulatively increase with each of the following three criteria met: if the household is ranked among the top 25% of all investors based on equity holdings at any point in time during the sample period; if the investor's portfolio is more concentrated than the median investor's; if the investor has ever traded options at least once during the entire sample period. The most sophisticated investors in Finland have a Sophistication score of 3, while the least sophisticated have a Sophistication score of 0. We divide investors into two sub-groups based on their sophistication. Group 1 includes the least sophisticated investors, whose sophistication scores are either 0 or 1. Group 2 includes the most sophisticated investors, whose sophistication scores are either 2 or 3. We then re-estimate our hazard model framework separately for Groups 1 and 2. Since we use sophistication dummies such as Option User Dummy, Log (Equity Portfolio Value), and Portfolio Concentration in constructing the two sub-groups, these covariates are not used as independent variables in the regressions. We also control for size, B/M, momentum, unrealized returns, and calendar year- and month-specific effects. Robust standard errors are adjusted as in Lin and Wei (1989). Ties are handled using the Efron procedure. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: Impact of Transaction Costs on Individual Traders' Holding Period Decisions in Finland, Hazard Analysis

	(1) AdjIlliq	(2) Zerofreq	(3) Spread (%)	(4) AdjIlliq	(5) AdjIlliq	(6) AdjIlliq
Illiquidity measure	0.984***	0.105***	0.986***	0.976***	0.979***	0.988***
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Size					1.000***	1.000 ***
					<.0001	<.0001
B/M					0.963***	0.996***
					<.0001	<.0001
Momentum					2.178***	1.008***
					<.0001	<.0001
Unrealized Returns					1.000***	1.000*
					<.0001	0.068
Age						0.996***
						<.0001
Male Dummy						1.341***
						<.0001
Option User Dummy						1.890***
						<.0001
Log (Equity Portfolio Value)						1.118***
						<.0001
Portfolio Concentration						4.106***
						<.0001
Firm stratification	No	No	No	Yes	No	No
Household stratification	No	No	No	Yes	Yes	No
Calendar year/month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,304,232	2,304,232	1,804,860	2,304,232	1,722,183	1,522,71
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Panel B: Impact of Sophistication on Attention to Transaction Costs in Finland		
Sophistication Group	(1)	(2)
Sophistication Score	0, 1	2, 3
AdjIlliq	0.992***	0.987***
	<.0001	<.0001
	Stock Characteristics	
Size	0.999***	0.999***
	<.0001	<.0001
B/M	0.925***	0.965***
	<.0001	<.0001
Momentum	2.087***	0.977
	<.0001	0.598
Unrealized Returns	1.000***	1.000***
	<.0001	<.0001
	Demographic Variables	
Age	0.996***	0.990***
	<.0001	<.0001
Male Dummy	1.374***	1.271***
	<.0001	<.0001
Firm stratification	No	No
Household stratification	No	No
Calendar year/month dummies	Yes	Yes
Number of Observations	809,296	395,442
Wald test	<.0001	<.0001

Appendix (Variable Definitions)

The Appendix describes in detail the variables used in the analyses.

Variable	Definition
Price ($)	Annual average of daily price
Market Cap ($mil)	Average market capitalization in millions of US dollars
B/M	Book-to-market ratio, calculated as the book value of equity at the end of previous year divided by market capitalization
Past Returns (-12, -2)	Cumulative returns during the past 10 months starting at month -12 and ending two months prior to a transaction
AdjIlliq	Adjusted Amihud ratio, calculated as the annual average daily ratio of absolute stock return to its dollar volume, then adjusted following Acharya and Pedersen (2005) to make it stationary
ZeroFreq	The percentage of zero-return days in a year, calculated following Lesmond, Ogden, and Trzcinka (1999)
Closing price Spread (%)	Calculated as the negative of the closing price divided by the purchase price minus one, following Barber and Odean (2000)
Commission (%)	Calculated as the commission charged by the brokerage for the trade scaled by purchase price
Effective Spread/Price (%)	The difference between the transaction price and the bid-ask midpoint multiplied by two and scaled by transaction price
Relative Spread/Price (%)	Quoted bid-ask spread divided by transaction price
Relative Spread/Mid (%)	Quoted bid-ask spread divided by the bid-ask midpoint
Depth	Midpoint of bid size and offer size in number of round lots; as depth tends to be skewed, we use log (1+depth) in our analyses
Unrealized Returns	Computed as (sale price - purchase price) / purchase price. For positions that are not closed at the end of the sample period, we assume the price on the last day of our sample period as the closing price.
Beta	Following Bali, Cakici, and Whitelaw (2011), beta is estimated every month using a regression of daily excess stock returns on daily excess market returns.
Ivol	Idiosyncratic volatility of each stock, calculated as the monthly standard deviation of its daily residual from Fama-French three-factor models

MaxPrc	Maximum daily return during the month prior to sales
Age in 1996 (1995)	Biological age of US (Finland) investor in 1996 (1995)
Married Dummy	Dummy variable that equals one for married investors
Male Dummy	Dummy variable that equals one for investors that are male
Professional Dummy	Dummy variable that equals one if the investor holds either a technical or managerial position
Retired Dummy	Dummy variable that equals one if the investor is already retired
Retirement Acct Dummy	Dummy variable that equals one if the transaction takes place in a retirement account, such as a 401(k)
Portfolio Concentration	Following Ivkovic, Sialm, and Weisbenner (2008), calculated as the sum of squared value weights of each stock in the investor's portfolio
Equity Portfolio Value ($ (€))	Total dollar (Euro) value of an US (Finland) investor's equity portfolio every month
Income ($K)	Self-reported income in thousands of dollars
Option User Dummy	Dummy variable that equals one if an investor has ever traded options during the entire sample period
Foreign Securities Dummy	Dummy variable that equals one if an investor has ever traded foreign assets, including ADRs, foreign stocks, or foreign mutual funds during the entire sample period
Short User Dummy	Dummy variable that equals one if an investor has ever shorted any security during the entire sample period