Do Individual Investors Ignore Transaction Costs?

Deniz Anginer
World Bank
Development Research Group

Snow Xue Han
College of Business
San Francisco State University

Çelim Yıldızhan
Terry College of Business
University of Georgia

December 2018

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 individual investors with longer holding periods choose to hold less liquid stocks in their portfolios, consistent with Amihud and Mendelson's (1986) theory of liquidity clienteles. The relationship between holding periods and transaction costs is stronger among more financially sophisticated households. Households whose holding periods are positively related to transaction costs also earn higher gross returns on their investments before accounting for transaction costs, suggesting that attention to non-salient transaction costs is an indication of investing ability. The main findings are confirmed 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 at least one particular type of non-salient cost, namely the cost of trading stocks, revealing a unique aspect of their rationality.

JEL Classifications: G11, G12, G14

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

---

* Deniz Anginer, World Bank, danginer@worldbank.org. Snow Xue Han, San Francisco State University, hanx@sfsu.edu. Çelim Yıldızhan, University of Georgia, celim@uga.edu. The authors wish to thank Jussi Keppo, Jeff Netter, Paolo Pasquariello, Bradley Paye, Annette Poulsen, Nejat Seyhun, Tyler Shumway, Noah Stoffman, Uday Rajan, seminar participants at the Ross School of Business at the University of Michigan, Terry College of Business at the University of Georgia, Erasmus Liquidity Conference, Tel Aviv University Finance Conference and the Western Finance Association meetings for comments and suggestions. All errors are our own. This paper is based on an earlier draft titled "Liquidity Clienteles."

1. Introduction

In a theoretical model, Amihud and Mendelson (1986) show that transaction costs cause a clientele effect in equity markets. Investors with longer holding periods hold stocks with higher transaction costs in equilibrium. Amihud and Mendelson (1986) emphasize that this positive relationship between transaction costs and holding periods represents a rational response in an efficient market where investors minimize their per-period transaction costs. Counter to Amihud and Mendelson's (1986) conjecture that investors understand and incorporate the impact of transaction costs in their investment decisions, recent 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 the trading records of households in the US and in Finland, to investigate whether individual investors are cognizant of one particular type of non-salient cost, namely cost of trading stocks, when making trading decisions. Specifically, we examine if individual investors hold illiquid securities with high transaction costs longer as stipulated by Amihud and Mendelson (1986), or whether they ignore transaction costs as suggested by the 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 but ignore hidden operating costs. Consistent with these findings, Gil-Bazo and Verdu (2008, 2009) document that there is a negative relation between mutual funds' before-fee performance and the fees they charge to investors. Surveys also suggest that retail investors do not understand all costs associated with investing in mutual funds (NASD Investor Literacy Survey 2003; Alexander, Jones, and Nigro, 1998).[1]

Besides mutual fund fees, there is also evidence that individuals do not pay attention to non-salient costs in other domains. Hossain and Morgan (2006), using a field experiment, show that buyers in Ebay auctions ignore shipping costs when the price of the item being auctioned is much higher than 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 little difference in 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.[2]

---

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

[2] Barber et al. (2009), using the complete transaction history of all investors in Taiwan, China, find 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.

However, losses incurred by individual investors after accounting for transaction costs do 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. Investors may trade when they experience income shocks (Lynch and Tan 2011) or when they experience exogenous liquidity shocks (Huang 2003). Second, even if most of the overtrading by individual investors could be attributed to overconfidence, there is no reason to claim that such investors do not pay attention to transaction costs. In this paper, we directly test whether investors pay attention to transaction costs by examining the relationship between transaction costs and the holding periods of individual investors. Rather than focus on trading performance alone, we analyze if individual investors understand the trade-offs between holding periods and transaction costs. This amounts to testing the first proposition in Amihud and Mendelson (1986) that investors will hold stocks with higher transaction costs longer compared to stocks with lower transaction costs.

In order to test Amihud and Mendelson's (1986) conjecture, 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. In particular, we use survival analyses and model investors' sell versus hold decisions at each point in time as a function of transaction costs using Cox hazard regressions.

Our findings are three-fold. First, we find that transaction costs are an important determinant of investors' holding periods after controlling for various household and stock characteristics. Specifically, we find that a stock in the highest transaction cost decile is 40% less likely to be sold than a stock that has lower transaction costs but with similar firm and investor characteristics, confirming Amihud and Mendelson's (1986) predictions. We check the validity of this finding by replicating our analyses in a 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 also 40% less likely to be sold compared to a stock that has lower transaction costs. Since the dataset from Finland includes 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 are robust to controlling for firm and household specific effects, additional controls and alternative measures of transaction costs.

Second, we find that households differ in how much attention they pay to the transaction costs of the securities they trade. Specifically, we find that investors who are more financially sophisticated pay more 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. In particular, we classify households who trade options, foreign securities and have short positions as financially more sophisticated. Our findings suggest that investor sophistication also plays a role in how much attention investors pay to transaction costs. Once again, we confirm our findings on financial sophistication using data from Finland.

Third, we find that investors who pay closer attention to transaction costs make better investment decisions overall. In particular, we find that households whose holding periods are positively correlated with transaction costs have higher *gross* returns (before accounting for transaction costs) than households whose holding periods are negatively correlated with transaction costs. This suggests that paying closer attention to transaction costs is a sign of investor ability. It would not be surprising to find that investors who do not pay attention to transaction costs have lower *net* returns (after accounting for transaction costs) than investors who do pay attention. As expected, we find that investors who pay attention to transaction costs realize lower spreads in trades.[3] These results are consistent with others who have shown significant variation in trading ability and performance in the cross-section of individual traders (see for instance, Grinblatt and Keloharju 2001 and Coval, Hirshleifer, and Shumway 2005).

There is likely to be endogeneity in the relationship between holding periods and measures of transaction costs used in this paper. As trading interest in a stock increases the costs associated with trading that stock decreases. We should note however the baseline or the average transaction costs of a given stock is likely to change slowly over time 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. Nonetheless, in order to address potential endogeneity concerns, we study investor behavior around periods of quasi-exogenous liquidity shocks.

Specifically, we examine how holding periods change around stock split events. A long line of literature documents a significant reduction in transaction costs and increase in liquidity subsequent to stock splits.[4] It has also been documented that post-split return performances of splitting firms are statistically indistinguishable from those of non-splitting firms with similar

---

[3] Spread is computed as in Barber and Odean (2000) as the percentage difference between the transaction price and the CRSP closing price.

[4] The prior literature suggests that liquidity increases after stock splits. For example, Schultz (2000) shows that the number of trades, especially the number of small trades, increases significantly subsequent to 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.

characteristics (Byun and Rozeff 2003). As such, splits do not appear to provide a signal to the market regarding the future prospects of the splitting firm as some theoretical papers have suggested. Taken together, these two findings would imply that the transaction costs channel is the main channel through which a stock split could affect the average holding period of households.

Consistent with the prior literature, we first verify that transaction costs decrease (stock liquidity increases) subsequent to stock splits in our sample period. Then we show that investors' average holding period declines in response to the increase in liquidity following stock splits. Our analysis suggests that the holding period for a stock decreases by about 34 trading days after a stock-split. This finding is economically significant as the average holding period for individual investors is 207 trading days.

The remainder of this paper is organized as follows. Section 2 develops the hypotheses tested in the paper. Section 3 describes the individual transaction 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 studies the implications of this relationship on individual investors' investment performance. Section 6 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 US results. Section 7 concludes.

2. Hypotheses

Amihud and Mendelson (1986) develop a model where investors with different exogenous holding periods trade securities with fixed transaction costs. They show that for investors with budget constraints, transaction costs result in a clientele effect, where investors with longer holding periods choose to hold illiquid stocks in equilibrium.

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 (Constantinides 1986, Vayanos 1998, Vayanos and Vila 1999, Heaton and Lucas 1996, Huang 2003, Lynch and Tan 2011, Lo, Mamaysky and Wang 2004). These dynamic models differ in their assessments regarding how transaction costs are priced, but they all generate the same conclusions as Amihud and Mendelson (1986) that investors' holding periods should correlate positively with transaction costs. Furthermore, while Amihud and Mendelson (1986) model each investor with a fixed holding period for all assets, recent dynamic models with endogenous trading horizons allow each investor to have different trading horizons for different assets. To accommodate the variation in investors' holding periods across different assets, we conduct our analyses at the transaction level. But we also perform and report our tests with household fixed effects, assuming that each household has a fixed baseline holding period. As both static and dynamic models predict that households' holding periods are positively related to transaction costs, our first hypothesis is:

H1: Holding periods of households 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. Goetzmann and Kumar (2008) also document that individual investors on average under-diversify their portfolio, and their portfolios underperform the market. 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 instance, Coval, Hirshleifer, and Shumway (2005) document strong persistence in the performance of individual investors' trades and show that some skillful individual investors are able to earn positive abnormal profits across different periods. Feng and Seasholes (2005) find that investors who are more sophisticated and possess more trading experience suffer less from the disposition effect bias. Ivkovic, Sialm and Weisbenner (2008) propose and empirically document that individual investors who hold more concentrated portfolios are financially more sophisticated and have better stock-picking skills that allow them to outperform other investors.

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 more 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. The first part of our second hypothesis is:

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

The correlation between holding periods and transaction costs is likely to impact a portfolio's *gross returns* (before accounting for transaction costs). We conjecture that paying closer attention to transaction costs could be a sign of market knowledge and trading ability. As previous studies have shown investor sophistication to be correlated with higher portfolio performance and lower levels of behavioral biases, we predict that households whose holding periods are positively correlated with transaction costs should earn higher *gross* returns compared to households whose holding periods are negatively correlated with transaction costs. Hence, the second part of our second hypothesis is:

> **H2.b**: Households who pay closer attention to transaction costs earn higher gross returns before accounting for transaction costs.

In testing the above hypothesis, we consider both raw portfolio returns as well as characteristics-adjusted returns computed as in Daniel et al (1997).

3. Data

3.1 Household Transactions and Demographics Information

This study uses two distinct datasets in order to explore the trading behaviors of households. The first dataset contains transactions for a subset of individual investors in the United States, while the second dataset contains more recent transactions of all investors in Finland. The individual trade data for the United States come from a major US discount brokerage house. It records the daily trades of 78,000 households from January 1991 to December 1996 and this is the same dataset used in Barber and Odean (2000).[5] A comparison of this dataset with Survey of Consumer Finances, IRS and TAQ data has shown it to be representative of US individual

[5] For a more detailed description of this dataset please refer to Barber and Odean (2000, 2001). We thank Terrence Odean for providing us with this dataset.

investors (Ivkovic, Sialm, and Weisbenner 2008, Ivkovic, Poterba, and Weisbenner 2005, and Barber, Odean, and Zhu 2006). 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 (both open-end and closed-end), American Depositary Receipts (ADRs), warrants, and options. About 66,000 of the 78,000 households trade common stocks, making about 800,000 transactions over the sample period. 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, such as income, age, gender, occupation and marital status.

To address concerns that our findings are specific to the data and sample period we employ, we repeat our analyses using another individual transactions 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- on a daily basis from January 1995 to December 2003.[6] 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, and it covers a more recent time period. For the purposes of this study, we ignore institutional trades and utilize only the trades of individual investors in Finland. Similar to the US dataset, the Finnish dataset reports for each transaction, the number of shares traded, the trading price, and the daily closing price. We can also observe the initial holdings for each account at the beginning of the sample period, which allows us to keep track of the holdings of households on a daily basis. While the dataset reports demographic information, such as age

[6] We thank Jussi Keppo for providing us with this dataset.

and gender for a subset of investors, it doesn't 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).

3.2 Measures of Transaction Costs

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 commonly used in previous papers that we are able to estimate for both the United States and Finland so as to make our results comparable across markets. Specifically, we use the adjusted Amihud ratio (*AdjIlliq*), Roll's measure (*Roll's C*), and zero return frequency (*Zerofreq*).[7]

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

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

$r_{i,d}$ 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

[7] For the US sample, we have also used transaction cost measures using high frequency intra-day tick data (from the TAQ database) and obtained similar results.

capitalizations of the market portfolio at the end of the month *t-1* and of the market portfolio in July 1962. The higher the adjusted Amihud ratio, the more illiquid the stock is.

The second measure we use is the Bayesian version of Roll's (1984) transaction cost measure, as estimated using Hasbrouck (2004) Gibbs sampler:

$$c_{i,t} = \begin{cases} \sqrt{-cov(r_{i,t}, r_{i,t-1})} & if\, cov(r_{i,t}, r_{i,t-1}) < 0; \\ 0 & otherwise. \end{cases} \tag{2}$$

It is based on the model $r_{i,t} = c_{i,t}\Delta q_{i,t} + \varepsilon_{i,t}$ where $q_{i,t}$ is a trade direction indicator, $c_{i,t}$ is the estimated half-spread that captures transaction costs, and $\varepsilon_{i,t}$ is an error term, all for stock *i* at time *t*.[8] The Bayesian estimation of this cost measure using the Gibbs sampler is described in detail in Hasbrouck (2009). The higher the Roll's C, the more illiquid the stock is.

The third measure uses the proportion of trading days with zero returns *(Zerofreq)* to capture transaction costs. The proportion of zero-return days as a proxy for liquidity was introduced by Lesmond, Ogden, and Trzcinka (1999) who argue that, on average, a zero return will be observed if transaction costs exceed the expected gains from trading. Following Lesmond et al. (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.

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

[8] From Roll's (1984) model, the half spread is equal to the square root of the negative serial correlation. When the sample serial correlation is positive, the estimator will be undefinable, and so we replace a default numerical value of zero following previous literature.

analyses. We realize the three measures used to proxy for transaction costs may not fully capture the real costs of trading and thus, we also repeat our analyses using realized transaction costs available in the data. In particular, we use bid-ask spreads and commissions for the US sample and utilize only bid-ask spread data for the Finland sample where commission information is unavailable.

Bid-ask spreads and commissions are calculated as in Barber and Odean (2000). We estimate spread using the negative of the transaction price divided by the closing price from CRSP minus one. Commission is the amount charged by the brokerage for the trade scaled by the dollar value of the trade. We use spread, commission, as well as the sum of the two as measures (*Spread+Commission %*) of actual transaction costs in our analyses. For brevity, we only report the results using the sum of spread and commission (*Spread+Commission %*) in Table 3. We obtain similar results using spread only or commission only in our analyses. We use spread only for the analyses in Finland and report the results in Table 8.

Table 1 reports summary statistics for stock and investor characteristics for the US.[9] Particularly, 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 very similar to those in the entire CRSP universe. For example, the median, 25th percentile and 75th percentile prices are the same for our sample of

[9] We report stock and investor characteristics for Finland in Table 7 and discuss the analyses for the Finland sample in Section 6.

stocks and 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 0.55 for the CRSP universe. The average and median size of our sample firms are also slightly larger than those of the CRSP universe. Overall, the differences are insignificant, indicating that our sample stocks are representative of the entire stock market during the sample period.

We observe that for our sample of stocks all transaction cost measures – adjusted Amihud ratio (*AdjIlliq*), *Roll's C*, and the proportion of zero return days (*Zerofreq*) - are positively skewed. The adjusted Amihud ratio has a mean of 5.04 and a median of 1.18, *Roll's C* has a mean of 1.91 and a median of 1.21 and *Zerofreq* has a mean of 7.14% and a median of 4.86%. The sum of spread and commission (*Spread+Commission %)* also exhibits positive skewness as the average *Spread+Commission %* is 2.12 while the median is 1.69. Similar patterns hold in the CRSP universe of stocks. The mean adjusted Amihud ratio is 5.56 for the CRSP universe, about 10% higher than that of our sample firms, suggesting that individual investors in our sample trade relatively more liquid firms. We observe similar patterns using *Roll's C* and *Zerofreq*.

Panel C reports the summary statistics for the US individual investors. Majority of the investors are in their 40s and 50s, with an average (median) age of 49.58 (48). 15% of the investors are retired. Only 10% of the primary US account holders for the transactions analyzed in this study are female, and 76% of the investors are married. 66% of the US individual investors in our transactions dataset 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. 14% of households have traded options, 22% have traded foreign securities, and 38% of the households 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*): *Holding periods are positively related to measures of transaction costs, after controlling for investor and stock characteristics.* We begin by computing a holding period for each transaction in the dataset.[10] 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).[11] This generates 799,469 holding period observations, with a median (mean) of 207 (550) trading days for retail investors in the United States. Since we examine a limited time period, there are transactions that are not closed by the end of the sample period. In the following section where we model holding periods in a hazard regression, we treat transactions that are not closed as censored observations.

[10] We also estimate portfolio level holding periods and transaction costs for each investor in the dataset. Our analysis of investor portfolios yields similar results as transaction-level analysis. As explained in the Introduction section, transaction-level analyses allow for the same investor to have different holding periods for different assets and as such are more in line with models that assume endogenous trading horizons, thus we report transaction-level analyses in the paper.

[11] We obtain similar results by alternatively defining the holding period as the number of trading days from the first purchase until the day when all outstanding positions are closed as in Feng and Seasholes (2005).

Figure 1 shows the median holding periods for US transactions grouped by investors' age, account type, the amount of capital the investor has invested in the stock market, as well the underlying stocks' transaction costs.[12] Investors who are older tend to have longer holding periods. Not surprisingly, the median holding period is also longer for stocks held in retirement accounts. Investors who have less capital invested in the market have longer holding periods suggesting that wealthier, more sophisticated investors trade more frequently. Most importantly, we observe that the average holding period for illiquid securities is significantly longer than the average holding period for the rest of the securities in our sample.

Next, we rank and assign the 799,469 holding period observations into ten groups based on their durations. For each group, we then calculate the average transaction cost, as well as price and market capitalization. The transaction cost measures are calculated as of the purchase day by averaging transaction costs over the previous 12 months. The results are reported in Table 2. There is a positive and largely monotonic relationship between holding period deciles and the average transaction costs of the stocks in each decile.[13] The adjusted Amihud illiquidity measure (*AdjIlliq*) increases from 0.92 for the shortest holding period decile to 1.84 for the longest holding period decile. Roll's C measure increases from 0.66 to 0.82, while the percentage of zero return days (*Zerofreq*) increases from 2.52% to 3.83% moving from the shortest holding period decile to the longest holding period decile. An even stronger result is observed for the sum of spread and

[12] In the figure, a stock is defined as Illiquid if it belongs to the highest transaction cost decile of stocks ranked according to the adjusted Amihud illiquidity ratio. *Other* category includes all other stocks not in the highest transaction cost decile. We get similar results using alternative transaction costs measures.

[13] In Table 2 we ignore the implications of censoring. For the purposes of only this table we assume that all transactions are closed by the end of the sample period. In all other analyses positions that are not closed by the end of the sample period are treated as censored observations.

commission (*Spread+Commission %*) as it monotonically increases from 1.08% of the transaction value for the shortest holding period decile to 2.77% of the transaction value for the longest holding period decile. The difference in the average illiquidity of the stocks traded in the longest and the shortest holding period deciles is significant regardless of the transaction cost measure employed. The univariate sort provides preliminary evidence that investors with longer (shorter) holding periods tend to trade stocks with higher (lower) transaction costs. As there is no clear pattern either for price (*Price*) or market capitalization (*Market Cap*) in holding period deciles, the monotonic relationship between holding periods and transaction costs does not appear to be a simple function of price or market capitalization.

Figure 2 depicts this relationship graphically. We sort stocks into two broad transaction cost groups based on the adjusted Amihud illiquidity measure. One group consists of stocks in the highest transaction cost decile while the other group holds the rest of the stocks in the other nine deciles. We plot Kaplan-Meier survival probabilities for these two broadly defined groups of stocks. 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 lighter line plots the survival probability of a representative stock in the highest transaction cost decile, while the darker line graphs the survival probability of a representative stock for the other nine deciles. The graph clearly shows that stocks ranked in the highest transaction cost decile have significantly higher survival probabilities than the rest of the stocks, suggesting that investors tend to hold stocks with higher transaction costs for longer periods of time before selling them. This again provides preliminary evidence that holding periods are strongly related to measures of transaction costs.

Next, 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.[14] Specifically, we model investors' sell versus hold decision using a Cox proportional hazard model with time-varying, as well as static explanatory variables.[15] The hazard model takes the following form:

$$h(t) = h_0(t)exp(\beta' X + \theta' Z_t) \quad (3)$$

This statistical framework describes how long an investor will hold a stock before selling it. 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 (such as gender). Z_t represents a vector of time-varying covariates which can take on different values at different points in time. $h_0(t)$ is called the baseline hazard rate and describes the average hazard rate when the independent covariates are equal to zero. Using the Cox (1972) estimator one can estimate coefficients on X and Z_t without specifying a baseline $h_0(t)$ 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' initial positions and trades, such as the total wealth

[14] 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.

[15] Details about estimating the proportional hazard model can be found in Cox and Oakes (1984).

invested in their portfolios, as well as dummy variables that take on a value of one if 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 which 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.[16]

Table 3 reports results from hazard regressions. Following standard reporting conventions, we report hazard ratios instead of 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 holding everything else equal. In the context of our model, each hazard ratio reflects the marginal effect of higher exposure to an explanatory variable on the probability of selling the stock.[17] A hazard ratio of less than one would suggest that, everything else held constant, a higher exposure to the explanatory variable would reduce the probability of selling the stock, hence increasing the likelihood of holding onto it. 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.

[16] Open stock positions, for instance, may be closed out in December for tax reasons.

[17] For example, a hazard ratio of 2 for the male dummy would suggest that the likelihood of a stock sale is twice more likely to happen per unit of time for male investors.

For brevity, we only report results using the adjusted Amihud illiquidity ratio for all specifications, while providing the baseline results for alternative measures of transaction costs.[18] Column (1) of Panel A shows that the estimated hazard ratio for the adjusted Amihud illiquidity ratio is 0.981 without controlling 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. More specifically, the average investor would be 9.3% less likely to sell a stock in the 75th percentile of illiquidity compared to a stock with median illiquidity which is equivalent to increasing the holding period by 23 trading days.[19] We obtain similar results using *Roll's C*, *Zerofreq*, and the *Spread+Commission %* in columns (2), (3), and (4) respectively. All estimated hazard ratios are less than one and statistically significant. The economic significance levels of these variables are also similar as well, regardless of the transaction cost measure we use. For example, the estimated hazard ratio for *Spread+Commissions %* reported in column (4) is 0.945, less than one and statistically significant. This suggests that a one standard deviation increase in *Spread+Commissions %* (2.95), would lead to a 15.4% reduction in the average household's sale likelihood, which is equivalent to increasing the holding period by 38 trading days.[20] These results are consistent with our univariate analysis in Table 2, confirming that investors' holding periods are longer for stocks with higher transaction costs.

---

[18] Results are qualitatively similar regardless of the transaction cost measure used and are available upon request.

[19] The adjusted Amihud illiquidity score is 1.18 for the median stock while it is 6.26 for the stock in the 75th percentile. Moving from the median stock to the 75th percentile stock would result in an increase of 5.08 in the adjusted Amihud illiquidity score. Since the hazard ratio for the adjusted Amihud measure is 0.981, an investor would be exp(ln(0.981)*5.08)=0.907 as likely to sell the stock in the 75th percentile of adjusted Amihud illiquidity as a stock with median adjusted Amihud illiquidity. This would make her 9.3% less likely to sell.

[20] A one standard deviation increase in *Spread+Commissions %* is equal to 2.95. Since the hazard ratio for the *Spread+Commissions %* measure is 0.945, an investor would be exp(ln(0.945)*2.95)=0.846 as likely to sell the stock.

As households could have different preferences and potentially have different holding periods, we control for the heterogeneity across households within the hazard framework. Specifically, 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 to a have 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 (5) of Table 3-Panel A we use firm and household level stratifications to account for household and firm fixed effects.[21] The estimated hazard ratio 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%, equivalent to increasing the holding period by 47 trading days.[22]

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 confirm that both household and

[21] We obtain similar results using Roll's C and the Zero-frequency measures, but for brevity we only report results the adjusted Amihud illiquidity measure. Results are available upon request.

[22] In unreported results, we conduct a similar analysis using dummy variables for the most illiquid and most liquid stocks where the dummy variable takes on a value of one if the stock belongs to the most illiquid (liquid) decile among all stocks in a given year. We find that a stock in the highest illiquidity (liquidity) group is 40% (20%) less (more) likely to be sold compared to a stock not belonging to that group.

stock fixed effects influence holding periods and verify that households do indeed differ in their baseline holding periods.

Considering heterogeneity across different stocks and investors, in columns (6) and (7) of Panel A in Table 3, we examine in detail how specific stock and investor characteristics affect households' trading decisions. The stock characteristics we control for in the model are firm size, *Size*; book-to-market ratio, *B/M*, and momentum, *Momentum*. Investor characteristics we include in the model are 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*.

Furthermore, we identify certain trader characteristics from each household's trading history, and define the following dummy variables: *Foreign Securities Dummy*, equals one if the household has ever traded foreign securities; *Option User Dummy*, set to 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*), as defined in Ivkovic et al. (2008) and computed as the sum of the squared value weights of each stock in a household's portfolio.

We only report results using the adjusted Amihud illiquidity ratio (*AdjIlliq*) for brevity, however, results using other transaction costs measures (including *Spread+Commissions %*) are very similar. We add stock characteristics first in column (6) of Table 3 Panel A, and then further control for investor characteristics in column (7). Since demographic information is only available for a subset of investors in the dataset, the number of observations is smaller in regression results reported in column (7). The basic finding is unchanged with these additional controls. The loading on the adjusted Amihud illiquidity measure in column (6) 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 would increase the probability of sale by 30.6%. The estimated hazard ratio for size is 0.649 and for the book-to-market ratio is 0.681, both of which are highly economically and statistically significant, suggesting that US individual investors tend to hold large and value stocks for longer periods. These hazard ratios would suggest that for each standard deviation increase in firm size investors are 10% less likely to sell the stock, and similarly for each standard deviation increase in book-to-market ratio they are 21.5% less likely to dispose of their holding.

For robustness, we also control for additional variables that prior studies have shown to affect individual investor trading decisions. In particular, prior studies have shown that individual traders tend to buy attention grabbing stocks and are also more likely to sell stocks that trade at prices below the purchase price (disposition effect).[23] To account for these behavioral biases we add

[23] The disposition effect is the tendency of individual investors to hold on to losing stocks for too long and to sell winners too quickly and has been shown to be a significant driver of trading behavior of individual investors (Grinblatt and Kellaharjou 2001). Moreover, 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 (2016) show that stocks with high conditional betas are also attention grabbing and attract many individual investors. In another related paper Kumar (2009) shows that investors tend to demand lottery-like stocks.

additional control variables in our main regression framework. These controls include a time varying dummy variable set to one if the price of a stock in an investor's portfolio is trading above its purchase price in a given day to capture the impact of the disposition effect as well as stock characteristics that are positively correlated with investor attention, namely idiosyncratic volatility (IVOL), maximum daily return over the past one month (MAX), as well as the CAPM beta of the stock (BETA). In unreported results, we find that the previously documented relationship between transaction costs and holding periods stays the same after including these additional controls.

Hazard ratios in column (7) for the adjusted Amihud measure as well as for stock-level characteristics are consistent with our results in column (6). In column (7), we also find that the hazard ratios on investors' trade related variables are all larger than one and statistically significant, implying that investors who have ever traded foreign securities, options, or have ever taken short positions tend to trade more frequently than investors that hold long positions in equities. Moreover, investors who have more money invested in the equity market and who hold more concentrated portfolios also trade more frequently. These findings collectively suggest that financially more sophisticated investors trade more frequently consistent with results in Figure 1. In the next section, we investigate in more detail the joint impact of transaction costs and investor characteristics on the decision to sell versus hold a security.

4.2 Impact of Investor Sophistication

In this section, we investigate the impact of heterogeneity across households on the relationship between transaction costs and holding periods of investors. In particular, we provide empirical evidence in support of the first part of our second hypothesis (*H2.a*): *The correlation between holding periods and transactions costs is higher for sophisticated investors.*

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. We use seven dummy variables to construct our sophistication measure. If the investor has income larger than $75K then the *High Income Dummy* is equal to one, if the investor works in a technical or managerial position then the *Professional Dummy* is equal to one, if the investor is ranked among the top 25% of all investors in terms of total equity holding then the *High Equity Holding Dummy* is equal to one, if the investor has ever traded an option then the *Option User Dummy* is one, if the investor has ever traded in foreign securities then the *Foreign Security Dummy* is equal to one, and if the investor has ever shorted any equity the *Short User Dummy* is equal to one. Ivkovic, Sialm and Weisbenner (2008) 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. In line with Ivkovic et al. (2008), we also assume that households with more concentrated portfolios are more likely to be financially sophisticated and assign a value of one to the *Concentrated Investor Dummy* when portfolio concentration is greater than 0.48, the median investor's level of portfolio concentration. Finally, we create a composite sophistication score (*Sophistication*) by adding up the seven dummy variables.[24] The value of *Sophistication* ranges from a minimum of 0 for the least sophisticated investor to a maximum of 7 for the most sophisticated investor.

[24] We use *High Equity Holding Dummy* and *Concentrated Investor Dummy* variables only in the construction of the Sophistication measure.

We re-estimate the baseline hazard regression (Model 1 in Table 3 Panel A) respectively for each group of investors according to their sophistication score. Group 0 corresponds to investors with a sophistication score of zero, while Group 7 includes investors with a sophistication score of seven. Figure 3 plots the hazard ratios of the transaction cost measure - *AdjIlliq* estimated from regressions performed for each group of investors based on their level of sophistication. X-axis indicates the value of the sophistication measure, and the y-axis shows the hazard ratio. The dotted lines show the 95th percentile confidence intervals of the estimated hazard ratios. The figure demonstrates that as investor sophistication increases, the hazard ratio on the adjusted Amihud illiquidity measure (*AdjIlliq*) decreases monotonically, suggesting that financially more sophisticated investors will hold stocks with higher transaction costs longer, providing evidence that more sophisticated investors pay closer attention to transaction costs.

Next, we include the interaction of the composite *Sophistication* measure with the transaction cost measures in our baseline model to examine how the relationship between transaction costs and holding periods differs among investors with different levels of sophistication. We report the results in Panel B of Table 3. Columns (1) and (2) of Panel B show that investors with higher *Sophistication* scores on average have shorter holding periods as in both columns the hazard ratios on *Sophistication* are greater than one and statistically significant. This is consistent with the results in column (7) of Panel A as the coefficients on *Foreign Securities Dummy, Option User Dummy, Short User Dummy, Log (Equity Portfolio Value)* and *Portfolio Concentration* are greater than one and statistically significant suggesting that the characteristics we associate with investor sophistication are on average correlated with shorter holding periods. While these results document that more sophisticated investors trade more frequently, it is the interaction of *Sophistication* with the transaction cost measures that indicates whether more sophisticated investors pay closer attention to transaction costs. If the relationship between holding periods and transaction costs is stronger for more sophisticated investors than for less sophisticated ones, the hazard ratio for the interaction term should be less than one.

Results in columns (1) and (2) in Panel B of Table 3 show that the hazard ratios for the interaction terms are indeed less than one and are statistically significant. Specifically, the hazard ratio of the interaction term *AdjIlliq* × *Sophistication*, is 0.995 before controlling for household specific effects in column (1) and it is 0.994 in column (2) after controlling for household specific effects with stratification. This implies that the relationship between holding periods and transaction costs is stronger for more sophisticated investors.[25] Hazard ratios in column (1) would suggest that some of the least sophisticated investors (i.e. *Sophistication*=1) are 12% less likely to sell when the adjusted Amihud ratio increases by one standard deviation, leading to a 28-day increase in the investor's holding period. In contrast, some of the most sophisticated investors (i.e. *Sophistication*=6) would be 27.1% less likely to sell for the same amount of increase in the adjusted Amihud ratio, resulting in an increase of 77 days in the investor's holding period. Results are stronger in column (2) when we control for household fixed effects using household stratification. More specifically, taking household fixed effects into account we predict that less sophisticated investors (*Sophistication*=1) will be 15.3% less likely to sell, while more sophisticated investors (*Sophistication*=6) will be 32.4% less likely to sell, leading to increases of 37 and 99 trading days

[25] The coefficients on the interactions of individual terms that constitute the Sophistication measure with AdjIlliq are not reported for brevity. The only interaction term that is marginally greater than one is AdjIlliq × Log (Income) which is statistically insignificant. All other interaction terms (AdjIlliq × Foreign Securities Dummy, AdjIlliq × Option User Dummy, AdjIlliq × Short User Dummy, AdjIlliq × Professional Dummy, AdjIlliq × Log (Equity Portfolio Value) and AdjIlliq × Portfolio Concentration) are statistically significant and less than one. Results are available upon request.

in holding periods, respectively. We obtain qualitatively similar results using alternative measures of transaction costs.

In summary, results in this section empirically verify our first (*H1*) and part of second (*H2.a*) hypotheses in line with Amihud and Mendelson's (1986) predictions. More specifically, we show that households choose to hold stocks with higher transaction costs for longer durations and that financially more sophisticated investors pay closer attention to the impact of transaction costs when they trade.[26]

5. Attention to Transaction Costs and Trading Ability

In this section, we study if paying closer attention to transaction costs is associated with overall investment ability. In particular, we hypothesize that a negative correlation between holding periods and transaction costs could indicate a lower level of financial sophistication, resulting in lower returns.

To identify households who pay more attention to transaction costs, we use the same hazard model as in Table 3, but instead of pooling across all households, we estimate the coefficient on the transaction costs variable (*AdjIlliq*) for each household individually. We use the correlation between holding periods and transaction costs for each investor as a proxy for how much attention the specific investor pays to transaction costs. In order to obtain robust estimates, we require that households make at least 50 round-trip trades over the sample period.[27] This leaves us with a sample of 2,192 households. For the majority of households in the dataset (over 60%), holding

[26] These findings are also in line with the predictions of dynamic models that endogenize the holding period decision as in Constantinides (1986), Vayanos (1998), Vayanos and Vila (1999), Heaton and Lucas (1996), Huang (2003), Lynch and Tan (2011), Lo, Mamaysky and Wang (2004).

[27] We obtain similar results using 20 or 30 trades instead of 50 trades.

periods are positively correlated with transaction costs, as the corresponding hazard ratio for the transaction costs variable, *AdjIlliq*, is less than one for these investors. This finding suggests that most individual investors in our dataset pay attention to transaction costs, as they hold illiquid securities in their portfolio for longer periods than they hold liquid securities.

We categorize households into two groups based on the magnitude of the hazard ratio on the transaction costs variable. In doing so, we classify investors as paying attention to transaction costs if their hazard ratio is less than one (Group 1 in Table 4) and not paying attention to transaction costs if the hazard ratio is greater than one (Group 2 in Table 4). We then compute 1-, 6- and 12- month holding-period raw returns for each transaction from the date of purchase,[28] as well as the 1-, 6- and 12- month characteristics-adjusted returns as in Daniel et al. (1997) to make sure that the differences in the returns are not driven by differences in stock characteristics. These results are reported in Table 4. Panel A reports returns for households whose hazard ratios are estimated with a 10% confidence level or higher, and Panel B reports results for all households.

There is a stark difference in the investment performances of these two groups. Households who pay attention to transaction costs (Group 1) earn 0.1%, 1.3%, and 2.9% more in raw returns and 0.1%, 1.0%, and 2.5% more in characteristics-adjusted excess returns than households who don't pay attention to transaction costs (Group 2) over 1-month, 6-month and 12-month holding periods respectively. The returns reported for all households in Panel B are similar in magnitude.

Table 4 also reports the average holding periods and the realized spreads for the two distinct groups of investors. *Spread %* is the ex-post realized bid-ask spread as described in 3.2. It is

[28] This is following Seru, Shumway and Stoffman (2010) and others. Using realized returns based on closed trades is problematic and can introduce biases. For instance, if the disposition effect is strong, only profitable trades would be closed.

calculated as the transaction price divided by the CRSP closing price minus one. This value is multiplied by minus one for purchases. The reported round trip spreads (*Spread %*) are computed as the sum of sale and purchase spreads. Investors who pay attention to transaction costs (Group 1) have lower amortized spreads of 0.675%. In contrast, investors who don't pay attention to transaction costs (Group 2) have higher round-trip spreads of 3.457%. The difference in spreads between the two groups of households is 2.8%, which is economically and statistically highly significant. This result indicates that investors who do not pay attention to transaction costs pay significantly higher transaction costs. This in turn suggests that the difference between the two groups' net returns after adjusting for transaction costs is even higher than the difference in their raw returns.

Overall, the findings in this section provide evidence in favor of our hypothesis *H2.b.* Investors who pay attention to transaction costs have on average better performance even before accounting for transaction costs, consistent with these investors being more sophisticated.

6. Robustness Tests

In this section, we conduct two additional analyses to show that our results are robust to potential endogeneity and selection concerns.

It is possible that our results are subject to an omitted-variables problem. If our transaction cost measures are related to certain unobserved variables which affect holding periods, then our results could be spurious. To address this concern, we use quasi-exogenous shocks to transaction costs. In particular, we use stock split events as shocks to transaction costs and examine investors' behavior around stock split events in section 6.1.

The second 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 test our main predictions. In order to address this criticism, we repeat our main analyses in section 6.2 utilizing another dataset which covers all Finnish individual investors' complete trading records from a more recent time period. Using an additional dataset from another country provides us with an "out-of-sample" test of our main findings.

6.1 Quasi-exogenous shocks to liquidity

In order to address the potential omitted-variables problem, we study investor behavior around periods of quasi-exogenous liquidity shocks. We consider stock splits as they are likely to affect holding periods through their effect on transaction costs.

A long line of literature documents a significant reduction in transaction costs and improved liquidity subsequent to stock splits (Schultz 2000, Desai, Nimalendran, and Venkataraman 1998, Kryzanowski and Zhang 1996, Conroy, Harris, and Benet 1990). Furthermore, the literature also documents that post-split performances of splitting firms are statistically indistinguishable from those of similar non-splitting firms (see for instance, Byun and Rozeff 2003), indicating that stock splits do not credibly signal improved performance. Taken together, these findings would suggest that if investors change their trading behavior subsequent to stock splits this should be attributable to the reduction in transaction costs rather than to a change in expected returns. Thus, we use stock splits as quasi-exogenous shocks and investigate investors' trading decisions around stock splits to provide more insights regarding the relationship between transaction costs and holding periods.

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. Following conventions in the stock split literature, we exclude 1,067 splits that distribute cash-dividends within a [-30, +30] days window around the split event to make sure that our results are not contaminated by other confounding effects. We also remove reverse splits and splits that have a split factor of less than 0.25. Our final sample includes 1,850 forward split events.

To examine how transaction costs change following stock splits, we regress the monthly transaction cost measure (*AdjIlliq*) on a time-period indicator, *After-Split dummy*, set to one if a month falls within the six-month, nine-month or twelve-month period subsequent to the split event, and zero otherwise.[29] We examine how transaction costs change within a certain period of time subsequent to stock splits, as we expect the effect of splits on transaction costs to decline over time. We also control for size, book-to-market and momentum. Table 5 presents the results. We find that the estimated coefficient on the After-Split dummy is always negative and statistically significant, suggesting that transaction costs, proxied by the adjusted Amihud Illiquidity ratio, decrease subsequent to stock splits. For example, in column (2), we observe that the coefficient on the After-Split dummy is -0.186 after controlling for stock characteristics. Considering that the median US stock has an adjusted Amihud illiquidity ratio of 1.36 (from Table 1 Panel B), the estimated -0.186 coefficient on the After-Split dummy suggests that transaction costs decrease by about 13% for the median stock after a split. Likewise, the estimated coefficient reported in column 4 (-0.248) suggests that transaction costs decrease by about 15% within nine months after a split.

[29] We also analyze two-month and three-month windows, and do not observe significant changes in stock liquidity. According to Lakonishok and Lev (1987), trading volume starts to increase from 7 months prior to the time a stock splits and then declines within 2 months subsequent to the split event. The pattern observed by Lakonishok and Lev (1987) might explain the insignificant results when we use two- or three-month windows.

If investors hold illiquid securities for longer periods as predicted by Amihud and Mendelson (1986), then the reduction in transaction costs subsequent to stock splits should lead to shorter holding periods. In other words, the likelihood of selling the splitting stock should increase following a split event. We examine individuals' trading behavior over the same 6-, 9- and 12-month periods subsequent to a splint event using a dynamic hazard regression framework.

To construct the appropriate dataset for the dynamic hazard framework, we split the duration period of the transaction into multiple periods. The first period covers the time frame from the purchase of the share until the split event. In this first period (pre-event), we assign a value of zero to the *After-Split dummy*. The second part is the time period from the split until the end of the event window of interest (i.e. we use three windows with the 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), for which the *After-Split dummy* takes on a value of zero.[30] Since it's possible for transactions to be open even after 6, 9, 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, *the After-Split dummy* will switch back to 0. In this framework, the *After-Split dummy* captures the marginal impact of stock splits on sale decisions over a distinct event horizon. The baseline hazard rate in the Cox regression model captures the increasing probability of a sale as time passes. *After-Split dummy*, therefore, captures the marginal impact of being in the split window period on the probability of a sale, and

[30] In the very rare instances where there are multiple splits before a transaction is closed, the After-Split dummy will be 1 during the post-split window but will switch back to 0 after each post-split event window.

does not simply capture a mechanical relationship due to the fact that probability of a sale increases as time passes on.

Table 6 reports the estimated results of the dynamic hazard regression. All the regression models in Table 6 control for calendar year and month specific effects, and models (2), (4), and (6) further control for stock characteristics - size, book-to-market and momentum. Given the reduction in transaction costs subsequent to stock splits, we would expect households to be more likely to reduce their holding periods, and thus we would expect the estimated hazard ratio on the *After-Split dummy* to be greater than one.[31] In all specifications, we find that the estimated hazard ratio for the *After-Split dummy* is greater than one and statistically significant at the 1% level. For example, the estimated hazard ratio for the *After-Split dummy* in model (2) is 1.183, indicating that investors are 18% more likely to sell a stock in the six months subsequent to its split, controlling for stock characteristics. The reduction in average holding period subsequent to a stock split in the US is equivalent to 32 trading days, which is economically highly significant since the median holding period is 207 days.

Our results are robust across different event windows. The hazard ratio on the *After-Split dummy* takes on a statistically significant value of 1.192 for the nine-month window analysis, and 1.198 for the 12-month window analysis after controlling for stock characteristics. These results suggest that investors are about 20% more likely to sell their stock holding within the first year subsequent to the split, which is roughly equivalent to reducing the representative investor's holding period by about 34 trading days. Using stock splits as quasi-exogenous shocks to

[31]The dynamic hazard regression method we apply naturally increases the number of observations. To efficiently estimate such a huge panel of data with likelihood function, we follow Allison and Christakis (2006).

transaction costs, we find results consistent with our prior findings as individual investors reduce their average holding periods significantly after a decline in transaction costs.

6.2 Out of Sample Test: Transaction Level Analyses in Finland

In this section, we provide another robustness analysis in support of our main findings. There may be sample selection concerns since the US transaction data is from the 1990s, and only covers a subset of investors. To address this concern, we replicate our analyses using transaction-level data from Finland, which covers individual investors' complete trading records between 1995 and 2003. Similar to the US dataset, the Finnish dataset reports for each transaction, the number of shares traded, the trading price, and the daily closing price. There are approximately two million transactions with available information from 303,235 households over the sample period. We can also observe the initial holdings for each account at the beginning of the dataset, allowing us to keep track of the total holdings of all households on a daily basis. For a subsample of investors, there is additional demographic information, such as age and gender. However, unlike for the US, we do not have information about income, occupation, and marital status. To calculate stock and firm characteristics, we obtain data from Datastream.

Table 7 reports summary statistics for stock and investor characteristics for Finland. 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. *Roll's C* has a mean of 2.96 and a median of 2.30 while *Zerofreq* has a mean of 21.90% and a median of 20.64%. Finally, we estimate the realized *Spread%* following Barber and Odean (2000), and find that it also has a positive skew.[32] The mean *Spread%* is 0.083 while the median is close to 0. Size is also positively skewed, with the average market capitalization approximately 10 times as large as the median one.

The average (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%. 4% of households have traded options at least once and less than one percent of the households in Finland 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.

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

Panel A of Table 8 reports results for the hazard model run on the transaction-level Finnish dataset. Results are remarkably similar to our findings in the US. In the baseline models, the hazard ratio of the adjusted Amihud illiquidity measure (*AdjIlliq*) in column (1) is 0.984, less than one and statistically significant. This indicates that if transaction costs (*AdjIlliq*) increase by one

[32] The commission data is not available for the Finnish dataset.

standard deviation (10.25) in Table 8, the investor is 15.2% less likely to sell that stock. As the median (mean) holding period for Finland investors are 100 (387) trading days, this decline in the likelihood of sales means that the representative investor's holding period will increase by an additional 18 trading days. We obtain similar results using *Roll's C* (column 2), *Zerofreq* (column 3), and the *Spread%* (column 4) measures. After we control for household and firm specific effects using stratification in column (5), the estimated hazard ratio on 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 the adjusted Amihud Illiquidity ratio, the representative investor is 22% less likely to sell compared to before, leading to an increase in the representative investor's holding period by 28 trading days.

To explore how stock, investor and trade characteristics affect holding periods, we include additional controls in the regressions reported in columns (6) and (7). Controlling for stock characteristics in addition to household specific effects in column (6) yields a less than one hazard ratio (0.982) on the adjusted Amihud illiquidity measure (*AdjIlliq*) which is statistically significant. Then, in column (7) we control for both investor and stock characteristics available in the dataset. The estimated hazard ratio for the adjusted Amihud illiquidity measure (*AdjIlliq*) is statistically significant at 0.982, suggesting that the average investor is 17% less likely to sell a stock when the stock's transaction cost increases by one standard deviation, leading to an increase in the representative investor's holding period by 20 trading days.

The hazard ratios for investor characteristics are also quite similar to those for the US. 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. 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 a fewer number of 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 prefer to hold smaller firms for longer periods. Altogether, results in Panel A of table 8 are very similar to our US findings reported in Table 3. Individual investors in Finland are also cognizant of and pay attention to transaction costs when they make trading decisions.

In Table 8 Panel B, we further investigate how household characteristics introduce 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. Investors who invest more capital in the stock market – higher *Log (Equity Portfolio Value)*, have experience trading options – one for *Option User Dummy*, or concentrate their investment in fewer number of securities – higher *Portfolio Concentration* are assumed to be financially more sophisticated.[33] Since the Finland transaction data doesn't provide information regarding investors' income, their professions, or whether the investor has ever traded

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

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. Column (1) of Panel B in Table 8 reports the hazard ratio of the Amihud illiquidity measure interacted with the overall sophistication measure, *AdjIlliq×Sophistication*. The interaction term is statistically significant and less than one (0.996). This result suggests that Finnish investors who are more financially sophisticated pay greater attention to transaction costs, and as such hold illiquid stocks for a longer period of time than less sophisticated investors do. In column (2), we further control for household specific effects using stratification and find similar results. Using alternative transaction cost measures, including *Roll's C*, the proportion of zero return days, (*Zerofreq*), as well as actual transaction costs, (*Spread%*), generates similar 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.

7. Conclusions

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. It provides a direct link between investors' holding periods and transaction costs and empirically verifies Amihud and Mendelson's (1986) theory of liquidity clienteles.

Three main conclusions follow from our analyses. First, we show that transaction costs are an important determinant of investment decisions of individual investors. Consistent with theoretical models of investor behavior, households rationally reduce the frequency with which they trade illiquid securities 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.

Third, we show that paying attention to transaction costs has important implications for investment performance. Households whose holding periods are negatively related to transaction costs earn lower gross returns compared to households whose holding periods are positively related to transaction costs, suggesting that investors who pay closer attention to transaction costs have better trading ability.

To address endogeneity and selection concerns, we study how investors' holding periods change around quasi-exogenous changes in transaction costs. We find that investors shorten their holding periods subsequent to stock split events when stock liquidity increases. We also extend our analyses using trading records from Finland, which include all trades made by every individual investor in Finland. We document that results in Finland are consistent with those in the US which suggests that our findings are unlikely to be driven by the specific sample of investors and the time period we study.

Our findings challenge the notion that individual investors ignore non-salient costs when making investment decisions. In this paper, using the trading records of households in the US and in Finland, we show that individual investors are cognizant of at least one particular type of non-salient cost, namely the cost of trading stocks, revealing a unique aspect of their rationality.

References

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

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

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,” 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, UC Davis

Bris, A., W. Goetzmann, and N. Zhu, 2007, “Efficiency and the Bear: Short Sales and Markets Around the World,” Journal of Finance, 62: 1029–1079

Byun, J. and M. S. Rozeff, 2003, “Long-run performance after stock splits: 1927 to 1996,” 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): 1145-77.

Choi, N., M. Fedenia, H. Skiba, and T. Sokolyk, 2017, “Portfolio concentration and performance of institutional investors worldwide,” Journal of Financial Economics, 123, 189-208

Christakis, N.A. and P. D. Allison, 2006, "Mortality after the Hospitalization of a Spouse," New England Journal of Medicine, 354, 719-730

Conroy, Robert M., Robert S. Harris, and Bruce A. Benet, 1990, "The effects of stock splits on bid-ask spreads," Journal of Finance 45, 1285-95

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

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

Desai, Anand 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 21,159-83

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," 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," 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?” Journal of Finance, 56, 589–616

Hasbrouck, J., 2009, “Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data,” Journal of Finance 64(3): 1445–1477

Hossain, T. and J. Morgan. 2006, “Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay,” B.E. Journals in Economic Analysis and Policy: Advances in Economic Analysis and Policy, 6(2): 1–27.

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

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

Ivkovic, Z., J. Poterba and S. J. Weisbenner, 2005, “Tax-Motivated Trading by Individual Investors,” American Economic Review, 95(5): 1605-1630

Lakonishok, Josef, and Baruch Lev, 1987, “Stock Splits and stock dividends: Why, who and when,” Journal of Finance 42, 913-932

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 55,429-50

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

Simutin, M., 2013, “Cash Holdings and Mutual Fund Performance,” Review of Finance, 18 (4): 1425-1464

Stoffman, N., 2014, “Who trades with whom? Individuals, institutions, and returns,” Journal of Financial Markets 21: 51-75

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

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

Table 1 reports 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 month starting at month -12 and ending two months prior. AdjIlliq is the adjusted Amihud illiquidity ratio. Roll's C is the Bayesian estimate of Roll's (1984) transactions cost measure. Zerofreq is the proportion of zero-return days which reports the percentage of zero-return days within a year. All liquidity measures are annual averages as defined in the text. 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. Commission is calculated as the commission paid divided by the value of the purchase. 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. 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 head of the household is a male. Professional Dummy is a dummy variable and is 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 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 reports the average total dollar value of a household's equity portfolio. 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 fund, 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	0.71	0.30	0.57	0.91	0.63
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
Roll's C	1.91	0.50	1.21	2.49	2.22
ZeroFreq	7.14%	0.00%	4.86%	10.42%	8.80%
Spread + Commission (%)	2.12	0.53	1.69	3.37	2.95
Panel B:	CRSP Stock Characteristics				
Price ($)	20.19	4.75	11.88	23.25	298.71
Market Cap ($M)	850.68	23.20	80.63	336.22	4057.39
B/M	0.72	0.30	0.56	0.91	0.63
Past Returns (-12, -2)	0.24	-0.10	0.08	0.29	2.12
AdjIlliq	5.56	0.38	1.36	7.55	7.92
Roll's C	2.00	0.52	1.27	2.60	2.33
ZeroFreq	7.53%	0.00%	4.86%	11.11%	9.25%

	Mean	P25	Median	P75	Std. Dev
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, univariate analysis

Table 2 presents univariate analysis of the relationship between stock illiquidity and the holding period for individuals in the US between 1991 and 1996. Holding Period is the number of trading days elapsed between the purchase and the first sale of that asset by a specific investor. Holding period for censored transactions, i.e. when sales of the assets are not observed, are estimated assuming the asset is sold on the last available date in the dataset. All transactions are ranked into ten portfolios based on the length of their holding periods. Averages for the various transaction cost measures and characteristics for the underlying securities are then calculated for each portfolio. All illiquidity measures are annual averages and are defined as previous in the text. Price is the annual average of daily closing prices. Market Cap is the average market capitalization in millions of dollars. AdjIlliq is the adjusted Amihud illiquidity ratio. Roll's C is the Bayesian estimate of the Roll's (1984) transaction cost measure. Zerofreq is zero return frequency which measures the proportion of zero-return days per year. 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 Commission is calculated as the commission paid divided by the value of the purchase. By definition, as the values of AdjIlliq, Roll's C and Zerofreq increase, liquidity decreases. High minus Low reports the differences for each variable between transactions groups with the longest and shortest holding-periods. Statistical significance at the 10%, 5% and 1% levels are denoted by *, **, and ***, respectively.

	Holding Period	Price $	Market Cap ($M)	AdjIlliq	Roll's C	Zerofreq (%)	Spread + Commission (%)
Low	6	33.46	7,283	0.9172	0.6608	2.52	1.0784
2	20	31.36	6,880	0.9865	0.6834	2.72	1.6776
3	44	30.25	7,153	1.0592	0.7054	2.84	1.9049
4	79	30.16	8,410	1.0828	0.7072	2.85	2.0700
5	127	33.46	8,601	1.2310	0.7337	3.06	2.1643
6	194	30.29	9,547	1.2279	0.7312	3.03	2.2466
7	294	30.51	9,814	1.3938	0.7382	3.09	2.3074
8	470	28.50	9,712	1.5231	0.7312	3.14	2.4547
9	771	30.49	11,508	1.6347	0.7425	3.40	2.5750
High	1225	34.37	9,686	1.8397	0.8182	3.83	2.7700
High-Low	1219***	0.91	2,403***	0.9224***	0.1575***	1.32***	1.6916***

Table 3: Impact of Transaction Costs on Households' Holding Periods in the US, hazard analysis

Table 3 examines the impact of transaction costs on individual investors' holding periods in the US using a hazard model framework. Panel A reports the estimated hazard ratios from the hazard regressions for US households between 1991 and 1996, where the conditional probability of sale is the dependent variable. Independent variables include the adjusted Amihud illiquidity ratio; firm characteristics; a set of demographic controls; trade variables; as well as the interactions of transactions costs measures with investor characteristics. Proxies for transactions costs (AdjIlliq, as well as Roll's C and Zerofreq) are calculated averages over the previous 12 months before each sale transaction. 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. Commission is calculated as the commission paid divided by the value of the purchase. Size is the market value of equity. 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. Age refers to the age of the head of the 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 the 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 average total equity holdings. 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. Panel B investigates if investor sophistication affects investors' attention to transaction costs. Investor sophistication is presumed to cumulatively increase with each of the following 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 seven (7), while the least sophisticated have a Sophistication score of zero (0). Calendar month dummies (not reported) are twelve dummy variables that equals 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. Wald test is used for each additional set of regressors. P-values are reported below each coefficient. Statistical significance at the 10%, 5% and 1% levels are denoted by *, **, and ***, respectively.

Panel A: Hazard Regressions, Impact of Transaction Costs on Individual Traders' Holding Period							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	AdjIlliq	Roll's C	Zerofreq	Spread + Commission (%)	AdjIlliq	AdjIlliq	AdjIlliq
Illiquidity measure	0.981***	0.982***	0.322***	0.945***	0.973***	0.981***	0.983***
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Stock characteristics							
Size						0.649***	0.724***
						<.0001	<.0001
B/M						0.681***	0.889***
						<.0001	<.0001
Momentum						1.135***	1.113***
						<.0001	<.0001
Demographic variables							
Age							0.997***
							<.0001
Log (Income)							0.927***
							<.0001
Married Dummy							0.959***
							<.0001
Male Dummy							1.103***
							<.0001
Professional Dummy							1.001
							0.891
Retirement Acct Dummy							0.912***
							<.0001
Retired Dummy							1.076***
							<.0001
Trade variables							
Foreign Securities Dummy							1.245***
							<.0001
Option User Dummy							1.395***
							<.0001
Short User Dummy							1.877***
							<.0001
Log (Equity Portfolio Value)							1.079***
							<.0001
Portfolio Concentration							3.228***
							<.0001
Firm stratification	No	No	No	No	Yes	No	No
Household stratification	No	No	No	No	Yes	Yes	No
Calendar month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	799,469	799,469	766,168	778,052	799,469	589,794	156,350
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Panel B: Hazard Regressions, Impact of Sophistication on Attention to Transaction Costs in US		
	(1)	(2)
	AdjIlliq	AdjIlliq
Illiquidity measure	0.988**	0.984**
	0.020	0.011
Age	0.993***	0.994***
	<.0001	<.0001
Married Dummy	0.879***	0.873***
	<.0001	<.0001
Male Dummy	1.046***	1.045***
	0.001	0.005
Retirement Acct Dummy	0.840***	0.862***
	<.0001	<.0001
Retired Dummy	1.379***	1.401***
	<.0001	<.0001
Illiquidity × Age	1.000***	1.000***
	<.0001	<.0001
Illiquidity × Married Dummy	1.003*	1.006***
	0.053	0.010
Illiquidity × Male Dummy	0.993**	0.992**
	0.021	0.043
Illiquidity × Retirement Acct Dummy	0.999	0.997
	0.441	0.251
Illiquidity × Retired Dummy	0.990***	0.988***
	<.0001	<.0001
Sophistication	1.265***	1.272***
	<.0001	<.0001
AdjIlliq × Sophistication	0.995***	0.994***
	<.0001	<.0001
Stock Controls	Yes	Yes
Household stratification	No	Yes
Calendar month dummies	Yes	Yes
Calendar year dummies	Yes	Yes
Number of Observations	156,350	156,350
Wald test	<.0001	<.0001

Table 4: Attention to Transaction Costs and Trading Ability

Table 4 reports returns for two groups of households based on the magnitude of their hazard ratios for the transaction cost variables. We estimated the hazard ratio of transaction cost measures for each US household using the hazard regressions described in Section 5. The estimated hazard ratio is used as a proxy for how much attention the individual pays to transaction costs. To get robust estimates, households are required to have made at least 50 trades during the sample period to be included in the analysis. Group 1 include investors whose hazard ratios are less than one, while group 2 is for investors whose hazard ratios are bigger than one. 1-, 6-, and 12-month returns are calculated starting from the date of purchase. Characteristics adjusted returns are calculated as returns net of characteristics matched portfolios, as in Daniel et al. (1997). Spread is estimated as in Barber and Odean (2000) and is equal to trading price divided by the CRSP closing price minus one for sales. The ratio is multiplied by minus one for purchases. Round trip spreads are the sum of sale and purchase spreads and reported below. Transactions with a purchase or sale price less than $2, and holding periods less than 2 days, are excluded from the sample. Panel A reports returns for the two groups where the hazard ratio of the transaction cost variable, AdjIlliq, is estimated with a minimum significance level of 10%. Panel B reports returns for the full sample. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: Observations with AdjIlliq Hazard Ratio at >10% Significance

	Group 1 Pays Attention to Transaction Costs	Group 2 Doesn't Pay Attention to Transaction Costs	Group 1 - Group 2
1 Month return	0.018	0.018	0.001
1 Month return Char. adj.	-0.001	-0.001	0.001
6 Month return	0.079	0.066	0.013***
6 Month return Char. adj.	-0.001	-0.002	0.010***
12 Month return	0.161	0.132	0.029***
12 Month return Char. adj.	-0.010	-0.035	0.025***
Spread (%)	0.675	3.457	-2.782***
Holding Period (days)	100.0	156.9	-56.9***

Panel B: All Observations

	Group 1 Pays Attention to Transaction Costs	Group 2 Doesn't Pay Attention to Transaction Costs	Group 1 - Group 2
1 Month return	0.018	0.017	0.001**
1 Month return Char. adj.	-0.001	-0.002	0.002**
6 Month return	0.079	0.070	0.009***
6 Month return Char. adj.	-0.010	-0.019	0.009***
12 Month return	0.162	0.146	0.016***
12 Month return Char. adj.	-0.009	-0.027	0.018***
Spread (%)	1.093	3.321	-2.228***
Holding Period (days)	115.6	147.4	-31.8***

Table 5: Change in Transaction costs around Stock Splits

Table 5 reports the changes in transaction costs for splitting stocks in the US around their ex-split dates. There are 1,850 forward stock splits during our sample period with a split factor larger than or equal to 0.25. We estimate an OLS regression of stock transaction costs on a time period indicator – After-Split dummy, controlling for size, book-to-market and momentum. The dependent variable is the monthly adjusted Amihud illiquidity ratio. Size, book-to-market and momentum are estimated monthly. We look at the changes in transaction costs in certain event windows, namely 6 months, 9 months and 12 months after stock splits. The time period indicator – After-Split dummy equals to one for splitting stocks in months that fall within the specified event window subsequent to the splits, otherwise it is zero. Each coefficient reported below for the After-Split dummy comes from a single OLS regression. For brevity, the coefficients for size, book-to-market and momentum are excluded. For each event window, we first report OLS results without firm controls and followed by results with firm controls in the adjacent columns. Standard errors are reported below each estimated coefficients. Statistical significance at the 10%, 5% and 1% levels are denoted by *, **, and ***, respectively.

Window	6-months		9-months		12-months	
	(1)	(2)	(3)	(4)	(5)	(6)
After-Split dummy	-0.193***	-0.186***	-0.250***	-0.248***	-0.293***	-0.289***
	(0.016)	(0.017)	(0.013)	(0.014)	(0.011)	(0.012)
Stock Controls	No	Yes	No	Yes	No	Yes
Observations	31,652	27,703	47,271	40,488	62,744	52,484
Adj. R^2	0.005	0.044	0.008	0.046	0.011	0.048

Table 6: Impact of US Stock Splits on Holding Period

Table 6 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. 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. The dependent variable is sale conditional on holding until day *t*. We employ three different event windows as defined before, namely 6-, 9-, and 12-month after the stock splits. The After-Split dummy is equal to 1 if that day falls into one of the specified event window. The table reports the estimated hazard ratios on the – After-Split- dummy. In all specifications, we control for time specific effects with calendar year and month dummies. We also control for size, book-to-market and momentum in models (2), (4) and (6). Stock characteristics are measured at the beginning of the month during which the sales are observed while year. For brevity, estimated hazard ratios for stock characteristics and calendar year and month dummies are not reported. P-values are reported below each coefficient. Statistical significance at the 10%, 5% and 1% levels are denoted by *, **, and ***, respectively.

Impact of Stock Splits on Holding Period Decisions

Window	6-months		9-months		12-months	
	(1)	(2)	(3)	(4)	(5)	(6)
After-Split dummy	1.344***	1.183**	1.328***	1.192***	1.310***	1.198***
	<.0001	<.0001	<.0001	<.0001	<.0001	0.0001
Stock controls	No	Yes	No	Yes	No	Yes
Calendar month dummies	Yes	Yes	Yes	Yes	Yes	Yes
Calendar year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	943,137	702,494	943,727	704,085	947,485	706,846
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

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

Table 7 reports descriptive statistics for stock and investor characteristics in Finland. Summary statistics are calculated by pooling annual observations over the 1996 - 2003 time period. We report the mean, median, standard deviation, as well as the 25th and 75th percentile values for all variables used in the study. All transaction costs measures are annual averages and are defined in the text. 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. Roll's C is the Bayesian estimate of Roll's (1984) transactions cost measure. 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.

Summary Statistics for Finland					
	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
Roll's C	2.96	0.66	2.30	3.88	3.51
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

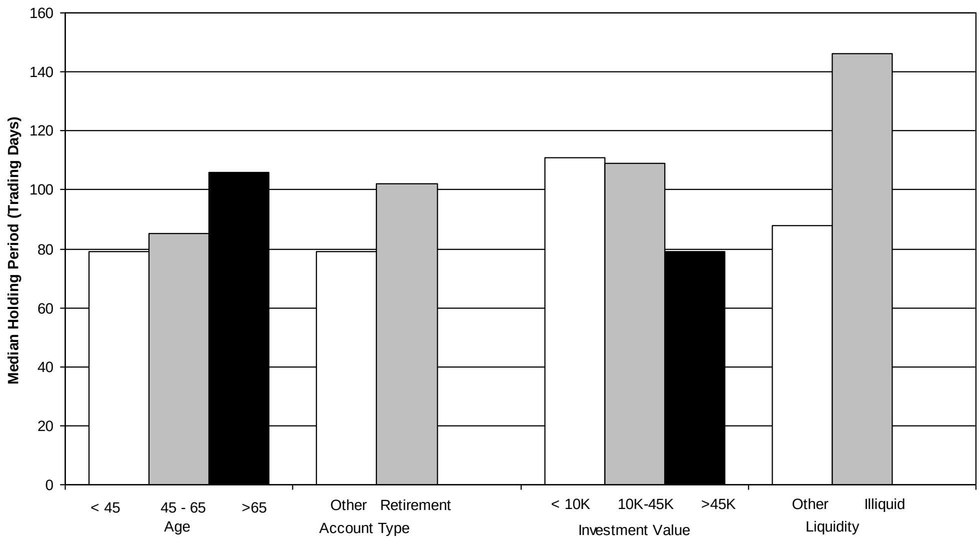
Table 8: Impact of Liquidity on Households' Holding Periods in Finland, hazard analysis

Table 8 examines the impact of stock liquidity on individual investors' holding periods in Finland using a hazard model framework. Panel A reports estimated hazard ratios from the hazard regressions where the conditional probability of sale is the dependent variable. Independent variables include transaction cost measures: the adjusted Amihud illiquidity ratio, (alternatively Roll's C, Zerofreq or Spread %); firm characteristics; a set of demographic controls; trade variables; as well as the interactions of our proxy for transactions costs with investor characteristics. Proxies for transactions costs (AdjIlliq, Roll's C and Zerofreq) are calculated over the previous 12 months prior to the sale transaction. 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. Size is the market value of equity. B/M is computed as the ratio of previous year-end book value to the most recent market capitalization. Momentum is the cumulative return over the period between month -12 to month -2. All stock characteristics are calculated at the beginning of the month that a sale takes place. Demographic variables include age and gender. Age is the biological age of the head of the household. Male Dummy is one if the head of the household is male. Trade variables for each investor are derived from all the transactions carried out by each specific investor in the dataset. Option User Dummy equals one if an investor has ever traded options at least once over the course of the sample period. Log (Equity Portfolio Value) is the logarithmic value of the household's total equity holdings. 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. Year Dummies are dummy variables that equal one if the sale transaction takes place during that particular year. Calendar month dummy is equal one if a sale takes place during that particular month. For brevity, estimated hazard ratios on the year and month dummy variables are not reported. 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 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 three (3), while the least sophisticated have a Sophistication score of zero (0). We also control for size, B/M, momentum, as well as 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: Hazard Regressions, Impact of Transaction Costs on Individual Traders' Holding Period Decisions in Finland							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	AdjIlliq	Roll's C	Zerofreq	Spread (%)	AdjIlliq	AdjIlliq	AdjIlliq
Illiquidity measure	0.984***	0.975***	0.105***	0.986***	0.976***	0.982***	0.982***
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
				Stock Characteristics			
Size						1.003***	1.001 ***
						<.0001	<.0001
B/M						0.993***	0.996***
						<.0001	<.0001
Momentum						1.014***	1.008***
						<.0001	<.0001
				Demographic Controls			
Age							0.994***
							<.0001
Male Dummy							1.290***
							<.0001
				Trade Variables			
Option User Dummy							1.713***
							<.0001
Log (Equity Portfolio Value)							1.157***
							<.0001
Portfolio Concentration							5.456***
							<.0001
Firm stratification	No	No	No	No	Yes	No	No
Household stratification	No	No	No	No	Yes	Yes	No
Calendar month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,304,232	2,304,232	2,304,232	1,804,860	2,304,232	2,131,366	1,522,716
Wald test	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Panel B: Hazard Regressions, Impact of Sophistication on Attention to Transaction Costs in Finland		
	(1) AdjIlliq	(2) AdjIlliq
Illiquidity measure	1.005***	1.013***
	<.0001	<.0001
Age	0.995***	0.995***
	<.0001	<.0001
Male Dummy	1.371***	1.339***
	<.0001	<.0001
Sophistication	1.588***	1.552***
	<.0001	<.0001
AdjIlliq × Age	1.000***	1.000
	<.0001	0.236
AdjIlliq × Male Dummy	0.990***	0.992***
	<.0001	<.0001
AdjIlliq × Sophistication	0.996***	0.997***
	<.0001	<.0001
Stock Controls	Yes	Yes
Household stratification	No	Yes
Calendar month dummies	Yes	Yes
Calendar year dummies	Yes	Yes
Observations	1,522,716	1,522,716
Wald test	<.0001	<.0001

Figure 1: Holding Periods of Households



This figure shows the median holding period for various investor and stock groups in the US. Age is the age of the investor. Account type denotes whether the account is a retirement account. Investment value is the average amount invested by the household in the stock market. A stock is defined as illiquid if it belongs to the highest decile of stocks ranked according to the adjusted Amihud illiquidity ratio. The holding period is calculated only for positions that are closed by the end of the sample period.

Figure 2: Survival Probabilities for Stocks in the United States

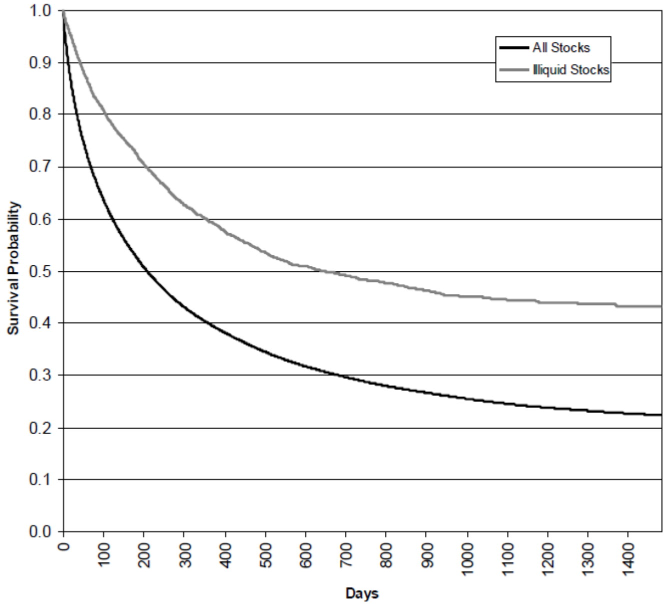


Figure 2 plots Kaplan-Meier survival probabilities for two groups of stocks held by households in the United States over the 1991-1996 time period. Illiquid stocks in the figure are stocks that belong to the top decile based on their adjusted Amihud illiquidity measure. The gray line stands for the probability of holding onto these illiquid stocks, and the black line stands for the probability of holding all the rest stocks.

Figure 3: Investor Sophistication and Hazard Ratios

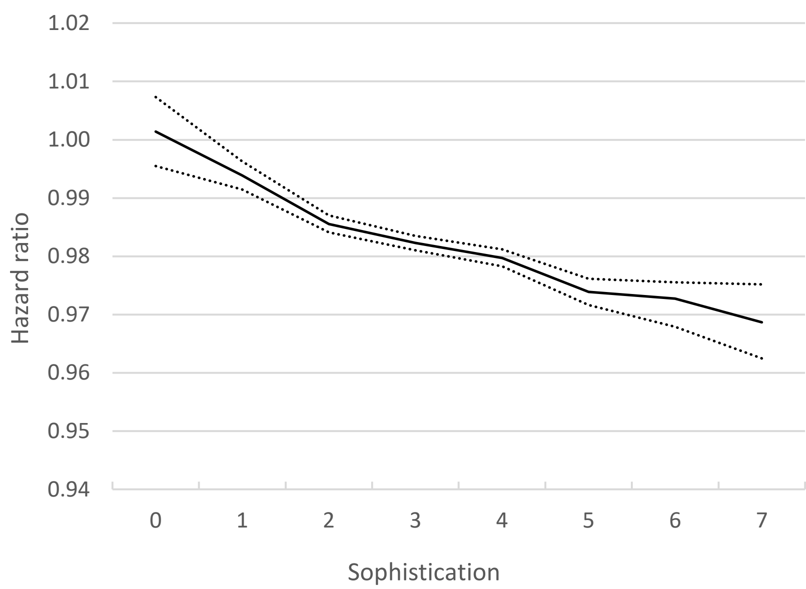


Figure 3 plots the hazard ratios of the variable *AdjIlliq* computed by running independent regressions for different groups of investors based on their level of sophistication indicated in the x-axis. Sophistication ranges from a minimum of zero (0) to a maximum of seven (7). The regression model used is the same as in column (1) of Panel B in Table 3. The dotted lines show the 95th percentile confidence intervals of the estimated hazard ratios.