Liquidity Clienteles: Transaction Costs and Investment Decisions of Individual Investors*

Deniz Anginer†

May, 2010

ABSTRACT

Theoretical papers link the liquidity premium to the optimal trading decisions of investors facing transaction costs. In particular, investors' holding periods determine how transaction costs are amortized and priced in asset returns. Using a unique dataset containing two million trades, this paper investigates the relationship between holding periods and transaction costs for 66,000 households from a large discount brokerage. I find that transaction costs are an important determinant of investors' holding periods, after controlling for household and stock characteristics. The relationship between holding periods and transaction costs is stronger among more sophisticated investors. Households with longer holding periods earn significantly higher returns after amortized transaction costs, and households that have holding periods that are positively related to transaction costs earn both higher gross and net returns. I show that there is correlation in the demand for liquid assets across households and, consistent with the notion of *flight to liquidity*, this demand increases during times of low market liquidity. Households with higher incomes and with higher wealth invested in the stock market supply liquidity when market liquidity is low.

Keywords: Transaction costs; Liquidity; Individual investor

JEL Classification Codes: G10, G11, G12

---

* I thank Yakov Amihud, Mario Macis, Paolo Pasquariello, Uday Rajan, Nejat Seyhun, Tyler Shumway, Jeffrey Smith, Noah Stoffman, Joe Warburton, Celim Yildizhan, and seminar participants at the University of Michigan and the World Bank Research Group for helpful comments. I thank Terrance Odean for providing the data used in this study. All errors are mine.
† Economist, World Bank Research Group. Please send correspondence to danginer@worldbank.org.

1. Introduction

Theoretical papers link the liquidity premium to the optimal trading decisions of investors facing transaction costs. An investor's required return on a stock subject to transaction costs will equal her required return in the absence of transaction costs plus these costs amortized over the investor's expected holding period. In a seminal paper, Amihud and Mendelson (1986) show that transaction costs cause a clientele effect, whereby investors with longer holding periods select to hold stocks with higher transaction costs in equilibrium. These *liquidity clienteles* drive how transaction costs are amortized and priced in asset returns. In theoretical models where the holding period is determined endogenously, the frequency with which investors trade illiquid securities subject to high transaction costs determine the holding period over which these transaction costs are amortized. If investors significantly reduce their trading of illiquid securities (Vayanos 1998, Constantinides 1988, Heaton and Lucas 1996) then amortized transaction costs will be low and investors will demand only a small liquidity premium to hold illiquid assets. If, on the other hand, investors have frequent trading needs because of income shocks (Lynch and Tan 2007), exogenous liquidity shocks (Huang 2003), or because they need to hedge non-traded risk exposure (Lo, Mamaysky and Wang 2004), then the resulting liquidity premium can be quite large.

Even though it is investors' trading decisions that provide the link between transaction costs and the liquidity premium on securities, lack of data on actual trades has made it difficult to empirically examine how investors behave in the presence of transaction costs. Using a unique dataset, this paper investigates the liquidity decisions of 66,000 households that made over two million trades using a large discount brokerage over a six-year time period. The focus of this paper is threefold. First, I examine empirically the relationship between investors' holding periods and the transaction costs of securities they trade and hold in their portfolios. Second, I investigate the impact of these liquidity decisions on investment performance. Finally, I examine the systematic decisions of households as a group over time. This paper differs from other empirical papers in this literature in that the focus is on investor (as opposed to stock) behavior.

I find that transaction costs play an important role in households' trading and investment decisions. Transaction costs are an important determinant of holding periods of investors after controlling for various household and stock characteristics. However, the effect of transaction costs on holding periods is much less than the effect predicted by the models of Vayanos (1998) and Constantinides (1988). The results in this paper offer an explanation for the discrepancy between the empirically-observed liquidity premium and the one predicted by these models in which the holding period is endogenously determined.[1] I find that households differ in how much attention they pay to the liquidity of the securities they trade and hold. More sophisticated investors tend to pay more attention to liquidity than less sophisticated investors. In addition, more sophisticated investors have holding periods that are strongly correlated with measures of transaction costs, while less sophisticated investors have negative correlations.

Household liquidity decisions have important implications for investment performance. I find that households with longer holding periods earn returns net of amortized transaction costs that are greater than the net returns of households with shorter holding periods. These results are consistent with Amihud and Mendelson (1986), who postulate that investors with longer holding periods earn rents that exceed amortized transaction costs for holding illiquid securities. This result drives the liquidity premium in their model. Consistent with the notion that sophisticated investors pay closer attention to liquidity, I find that households whose holding periods are negatively correlated with transaction costs earn lower *gross* and *net* returns. That is, households that do not pay attention to liquidity earn lower returns on both a gross and net basis.

I also find that there is systematic variation in the demand for liquid assets across households. Consistent with the notion of *flight to liquidity*, the demand for liquid assets goes up during times of low aggregate market liquidity, with households tending to buy liquid securities and sell illiquid securities. However, a subset of investors with deep pockets, i.e., those with higher incomes and higher levels of wealth, buys illiquid securities when there is a negative liquidity shock and, consequently, earns a premium in the process.

While investor decision making in the presence of transaction costs is important to better understand how liquidity is priced in the financial markets, it also has implications

[1] For empirical studies, see, for instance, Pastor and Stambaugh (2003), Acharya and Pedersen (2005), Amihud (2002), and Amihud and Mendelson (1986).

for investor welfare and public policy. This paper shows that expected holding periods and amortized transaction costs strongly impact the performance of household portfolios. One implication is that investment advisors should consider the expected holding period of investors when recommending illiquid stocks to their clients. The results in this paper also have implications for the efficacy of a securities transaction tax. Such a tax has been proposed to reduce excess speculation in order to reduce volatility and the influence of short-term investors on management (Stiglitz 1989, Tobin 1984, Summers and Summers 1990). This paper provides an empirical link between the magnitude of such a tax and its impact on trading frequency of retail traders.

This paper is also related to investor rationality and, in particular, to the increasingly popular notion that individual investors overtrade, losing substantial amounts to trading costs without any gain in performance.[2] Usually a behavioral bias, such as overconfidence, is proposed as an explanation for excessive trading by individual investors who tend to ignore transaction costs. Barber, Odean and Zheng (2005), for instance, show that investors pay attention only to the salient costs of mutual funds, but ignore hidden operating costs. The findings in this paper suggest that most investors are, to a large extent, cognizant of transaction costs when making trading decisions. The findings suggest that, as investors trade more frequently, they pay greater attention to the liquidity of the underlying stocks traded. A number of papers also document that a subset of retail investors displays greater financial sophistication and market understanding than the average retain investor, enabling them to earn positive abnormal returns.[3] In this paper, I show that sophisticated households are more likely to hold illiquid stocks over a longer time period and earn greater net returns as a result.

In a related paper, Atkins and Dyl (1997) study the relationship between turnover and bid-ask spreads for Nasdaq and NYSE stocks. They find a positive relationship between bid-ask spreads and holding periods, which they proxy with turnover. There are, however, two problems with using aggregate turnover to proxy for holding periods. First, aggregate turnover is an average across many investors and can be highly skewed in a

[2] Barber and Odean (2000) show that investors similarly ranked in terms of portfolio turnover have similar gross returns, but substantially different net returns after accounting for transaction costs. Barber et al. (2008), using a complete transaction history of all investors in Taiwan, find that individual investor losses equal 2.2 % of GDP, and that such loses are mainly due to transaction costs.

[3] See the discussion in Section 2.

market where a handful of investors trade to provide liquidity. Second, and more importantly, holding periods are based on trading decisions of investors who, ex-ante, consider the transaction costs of the underlying securities they trade. In a concurrent paper, Naes and Odegaard (2008) use transaction-level Norwegian data to show that turnover is indeed a poor proxy for actual holding periods of investors. Their focus is on asset pricing, and they show that turnover is priced in size-sorted portfolios while average holding period is not.

The remainder of the paper is organized as follows. The next section describes the empirical questions pursued in this paper. Section 3 describes the liquidity measures and the individual trade data used herein. Sections 4 to 6 present and discuss the main findings, and Section 7 concludes.

2. Hypotheses and Related Literature

Although empirical studies document that the effects of transaction costs on asset prices are both statistically and economically significant, there is a debate in the theoretical literature as to the direction and the magnitude of this relationship.[4] The debate centers on how investors make optimal trading decisions in the presence of transaction costs. The basic premise that the rate of return on a security should incorporate transaction costs is straightforward and uncontroversial. An investor who buys a security and expects to pay transaction costs when selling it will take this into account in valuing that security. An investor’s required return on a stock will equal her required return in the absence of transaction costs plus these costs amortized over the investor’s expected holding period. The liquidity premium required by investors to hold illiquid securities thus depends strongly on investors’ holding periods. The theoretical debate over the effect of transaction costs on asset prices arises primarily from differences in how investors’ holding periods are modeled.

One of the earlier papers to incorporate investors’ holding periods into asset pricing with market frictions is Amihud and Mendelson (1986). They develop a model where

[4] For empirical studies, see, for instance, Pastor and Stambaugh (2003), Acharya and Pedersen (2005), Amihud (2002), Chordia et al. (2000, 2001), Hasbrouck and Seppi (2001), and Huberman and Halka (1999).

risk neutral investors with different exogenous holding periods and limited capital trade securities subject to fixed transaction costs. Amihud and Mendelson show that transaction costs result in a clientele effect, with investors who have longer holding periods selecting to hold illiquid stocks in equilibrium. Amortized transaction costs of investors in each *liquidity clientele* group determine the liquidity premia for illiquid securities.

The static model with exogenous holding periods has been extended to incorporate dynamic decisions of investors. In models where the holding period decision is determined endogenously (Constantinides 1986, Vayanos 1998, Vayanos and Vila 1999, Heaton and Lucas 1996), the resulting liquidity premium is much lower. In these models, the marginal utility from trading is low and investors respond to transaction costs by turning over their portfolio less frequently. These models predict a liquidity premium on asset prices that is a magnitude smaller than transaction costs, but they also predict unrealistically low levels of trading activity and volume. In models where investors are forced to trade frequently (Huang 2003, Lynch and Tan 2007, Lo, Mamaysky and Wang 2006) the resulting liquidity premium can be large.

In all these models, the magnitude of the relationship between holding periods and transaction costs determines the liquidity premium in the market.[5] Using individual trade data, I test for the relationship between holding periods and transaction costs after controlling for a number of investor and stock characteristics. I also analyze the magnitude of the impact of transaction costs on holding periods, and compare the results to calibrated values in the models of Vayanos (1998), Constantinides (1986) and Lo, Mamaysky and Wang (2005). The first hypothesis is thus:

H1a: *Holding periods are positively related to measures of fixed 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-

[5] Although in this study I only focus on a subset of investors in the market, namely retail investors, a number of papers have shown that correlated trading by retail investors impact returns (Kumar and Lee 2006, Barber, Odean and Zhu 2006, and Hvidkjaer 2008).

perform their sales by a significant margin.[6] However, other studies have concluded 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, suggesting that some skillful individual investors might be able to earn positive abnormal profits. Using the same dataset as this paper, Goetzmann and Kumar (2008) find that the level of portfolio diversification is related to investor sophistication. Feng and Seasholes (2005) find that investor sophistication reduces a well known behavioral bias, the disposition effect. Given that previous studies have documented heterogeneity in the performance and investment decisions of individual investors, we should expect similar cross-sectional differences in the correlation between holding periods and transaction costs across investors in the dataset. Furthermore, we should expect this correlation to increase with investor sophistication and experience:

H1b: *The correlation between holding periods and transactions costs is higher for sophisticated investors.*

The second empirical question I address in this paper is how holding periods and transaction costs impact investment performance. In the Amihud and Mendelson (1986) model, it is the rents earned by investors with longer holding periods that drive the liquidity premium. Security prices reflect the marginal investor's holding period, and have to fall by the present value of transaction costs to induce the marginal investor to buy the security.[7] The price for the security with the lowest transaction cost, for instance, is set such that the investor with the shortest holding period is indifferent between investing in that security and the one with no transaction costs. Investors with longer holding periods earn a premium (rent) when investing in that security because their amortized transaction costs are lower, which implies:

[6] Barber and Odean (2000, 2001), using the same dataset as this paper, further show that investors lose substantial amounts to trading costs without any additional gain in performance, consistent with the hypothesis that individual investors are overconfident and tend to trade excessively.
[7] Vayanos and Vila (1997) show a similar result when securities are identical except for transaction costs.

H2a: *Investors with longer holding periods earn returns net of amortized transaction costs that exceed net returns of investors with shorter holding periods.*

The correlation between holding periods and transaction costs is likely to impact portfolio performance on both a gross and a net basis. Households that do not pay attention to transaction costs when they trade are likely to have lower net returns due to transaction costs. As mentioned earlier, previous studies have shown investor sophistication to be correlated with higher portfolio performance and lower levels of behavioral biases. A negative correlation between holding periods and transaction costs could, therefore, also indicate lack of financial sophistication and market knowledge, which is associated with lower gross returns. Consequently:

H2b: *Investors whose holding periods are negatively related to transaction costs earn lower gross and net returns.*

In other words, we would expect investors who do not pay attention to liquidity to make other trading mistakes, which result in lower gross returns.

Previous studies have shown that there is a common time varying component to liquidity across stocks (Chordia et al 2000, Hasbrouck and Seppi 2001, and Huberman and Halka 2001). Other studies have shown that this common component is priced in stock returns (Pastor and Stambaugh 2003, Acharya and Pedersen 2005, Korajczyk and Sadka 2008). It is not clear, however, what causes this common variation. Commonality in liquidity can arise from the supply side, if there is systematic variation in the costs of providing liquidity.[8] Commonality can also arise from the demand side, if a common factor such as volatility or uncertainty causes a systematic variation in the demand for liquidity.[9] Even with constant exogenous transaction costs, a time-varying liquidity premia can arise as investors' willingness to bear these costs changes over time. Vayanos (2004), for instance, develops a model with fixed transaction costs in which changes in

[8] Chordia, Roll and Subrahmanyam (2000) find some evidence of asymmetric information and inventory risk affecting the common component of liquidity. Comerton-Forde et al (2008) and Coughenour and Saad (2004), examining liquidity of stocks at NYSE overseen by the same specialist, provide some support for the supply side view. Huberman and Halka (2001), on the other hand, after failing to find inventory cost or asymmetric information based explanations for the systemic component of liquidity, conjecture that commonality emerges due to noise traders.

[9] Chordia et al. 2001 shows that trading activity covaries with liquidity.

market volatility affect systematic liquidity by creating correlated trading patterns among investors. By examining the actual trades of investors, I can test whether there is systematic variation in the demand for liquid assets and whether liquidity shocks apply (or transmitted) systematically across investors that can potentially cause market-wide effects:

H3a: *There is systematic variation in households' trades of illiquid stocks.*

If there is systematic variation in demand for liquid assets across investors, it is important to examine how this systematic demand varies over time with changes in aggregate level of market liquidity. If investors demand liquid securities at the same time when aggregate liquidity is low, the liquidity premium required to hold illiquid securities would be high. The literature, to a large extent, treats individual investors as noise traders who provide constant liquidity to the market. Kaniel, Saar, and Titman (2006), Campbell, Ramadorai, and Schwartz (2007), Stoffman (2008), and Griffin et al. (2003), investigating institutional and retail trades, provide evidence consistent with the notion that retail traders provide liquidity to meet institutional demand for immediacy. These studies, however, investigate short-term returns to institutional and individual buy/sell imbalances, and do not consider the liquidity level of the market or the liquidity level of the individual securities that are traded.[10] With individual trade data, I can examine the liquidity level of the securities bought and sold by individual investors, and examine whether there is a *flight to liquidity* among households. With individual trade data, I can also test if households are net demanders or suppliers of liquid securities when aggregate market liquidity is low:

H3b: *Households are net buyers of illiquid stocks when the market level of liquidity is low.*

There are likely to be cross-sectional differences in trading patterns in response to aggregate liquidity shocks. Investors with deep pockets can take advantage of investment opportunities during turbulent markets. The recent Goldman Sachs' agreement to sell $5

[10] In most of these studies, investors cannot be identified and their transactions cannot be tracked over time.

billion of perpetual preferred stock to Berkshire Hathaway illustrates both the adverse effects that market participants can create when seeking liquidity at the same time and the important role that external investors with deep pockets can play in providing liquidity. We can expect households with higher wealth/income levels to buy illiquid assets that have dropped in price:

H3c: *Households with higher income and wealth levels are net buyers of illiquid stocks when aggregate market liquidity is low.*

3. Individual Trade Data and Liquidity Measures

The main dataset for this paper comes from a major U.S. discount brokerage house and includes the daily trading records of 78,000 households from January 1991 to December 1996. These households hold a total of 158,034 accounts of various types including cash, margin, IRA and Keogh. In this study, I focus on the common stock investments of the households, which constitute nearly two-thirds of the total value of their investments in the dataset. About 66,000 of the 78,000 households trade common stock, making close to two million trades over the sample period. The transaction record includes number of shares traded, price and any commissions paid. The dataset also includes each household's month-end positions including the value of security holdings at market close on the statement date. For a sub-sample of households, the dataset includes demographic information, such as income, age, gender, occupation and marital status. A more detailed explanation of the dataset can be found in Barber and Odean (2000, 2001). A comparison of this dataset with Survey of Consumer Finances, IRS and TAQ data has shown it to be representative of U.S. individual investors (Ivkovic, Sialm, and Weisbenner 2006, Ivkovic, Poterba, and Weisbenner 2005, and Barber, Odean, and Zhu 2006).

Liquidity is a multi-faceted concept, and is usually defined in terms of the costs and risks associated with transacting financial securities. These costs relate to exogenous costs of transacting including price impact, asymmetric information and inventory risk. Given the multi- faceted and unobservable nature of liquidity, I use a number of different measures that have been previously utilized in the literature. The first is a Bayesian version of the Roll (1984) transaction cost measure:

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

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 transaction cost measure and $\varepsilon_{i,t}$ is an error term for stock *i* at time *t*. Equation (1) can be derived under the assumption that buyer- and seller-initiated trades are equally likely. The Bayesian estimation of this cost measure using the Gibbs sampler is described in detail in Hasbrouck (2006).[11]

The second measure is the Amihud illiquidity ratio, calculated as:

$$Illiq_{i,t} = \frac{1}{D_{i,t}} \sum_{d=1}^{D_{i,t}} \frac{\left| r_{i,d} \right|}{dvol_{i,d}} \tag{2}$$

where $D_{i,t}$ is the number of days in month *t* for stock *i*, $dvol_{i,t}$ is the dollar volume in day *d*, and $r_{i,d}$ is the daily return. While the bid-ask spread captures the cost of executing a small trade, the *Illiq* variable is akin to Kyle's lambda and is meant to capture the price impact of a trade. I adjust this measure as in Acharya and Pedersen (2005) to make it stationary and to remove outliers:

$$AdjIlliq_{i,t} = \min\ 0.25 + 0.30 \times Illiq_{i,t} \times M_{t-1}, 30 \tag{3}$$

where M_{t-1} is the ratio of the value-weighted market portfolio at the end of the month *t*-1 to that of the market portfolio in July of 1962.

The third measure used in this paper is the Pastor and Stambaugh (2003) reversal gamma:

[11] The Gibbs estimate is obtained from Joel Hasbrouck's website: http://pages.stern.nyu.edu/~jhasbrou/Research/GibbsEstimates2006/Liquidityestimates2006.htm.

$$r^e_{i,d+1,t} = \theta_{i,t} + \phi_{i,t} r_{i,d+1,t} + \gamma_{i,t}\mathrm{sign}(r^e_{i,d,t}) v_{i,d,t} + \varepsilon_{i,d,t} \quad (4)$$

Above, $r^e_{i,d+1,t}$ is the return in excess of the market return and $v_{i,d,t}$ is the volume on day *d* in month *t* for stock *i*. This measure is motivated by the Campbell, Grossman, and Wang (1993) model and is meant to capture temporary price fluctuations arising from order flow.

I also include in the analyses quoted and effective spread and quoted depth calculated from intra-day data. I use a 5-second delay to match trades with quotes and apply the same filters discussed in Hvidkjaer (2006). The quoted percentage spread is calculated for each trade as the ratio of the quoted bid-ask spread to the prevailing transaction price. The effective percentage half-spread is calculated for each transaction as the absolute value of the difference between the transaction price and the quote midpoint, divided by the bid-ask midpoint. The quoted depth is the average of quoted bid-ask lots multiplied by bid-ask quotes. In addition, I compute a *realized spread*, which is the ex-post realized bid-ask spread paid by the investors for each transaction in the dataset. The calculation is the same as in Barber and Odean (2000):

$$\begin{aligned} \mathrm{SprBuy} &= \frac{P_{crsp}}{P_{buy}} - 1 \\ \mathrm{SprSell} &= 1 - \frac{P_{crsp}}{P_{sell}} \end{aligned} \quad (5)$$

where P_{crsp} is the closing price from CRSP, and P_{sell} and P_{buy} are the purchase and sale prices from the dataset. This measure includes the bid-ask spread, market impact of the trade as well as the intra-day return on the day of the trade. The total spread is the sum of the realized buy and sell spreads. Previous studies (Korajczyk and Sadka 2008, and Eckbo and Norli 2002) have shown that there is high correlation among these liquidity measures and that there is a common component that accounts for most of the variation across individual liquidity measures.

There is likely to be endogeneity in the relationship between holding periods and liquidity measures used in this paper. As trading interest in a stock increases so does its liquidity. But we can also think of a stock as having a baseline exogenous cost component along the lines of Amihud and Mendelson (1986). Although the liquidity level of a penny stock, for instance, will increase with increased trading interest, it will not achieve the same level of liquidity of a large cap stock purely based on that increase.[12] Figure 1 illustrates this notion graphically. I plot the adjusted Amihud illiquidity ratio for IBM and Crown Petroleum Corp. over the 1991 to 1996 period. Although there is variation over time in the liquidity levels for both stocks, the average *AdjIlliq* ratio is significantly lower for IBM over the sample period. To capture this baseline component, I use annual averages of the liquidity measures in analyzing household holding periods. I later extend the analyses to incorporate time series variation in Section 6. Table 1 reports the summary statistics and correlations for the liquidity measures for stocks traded by households in the dataset.

4. Holding Periods and Transaction Costs

4.1. Transaction Level Analyses

To examine the relationship between holding periods and transaction costs, I first calculate a holding period for each transaction in the dataset. The holding period is defined as the number of trading days from the first purchase of a stock to the first sale.[13] This method provides 806,404 holding period observations. The average and the median holding period are 185 and 86 trading days respectively. Figure 2 shows the median holding periods for transactions grouped by investors' age, account type, the amount of capital they have invested in the stock market, as well as transactions grouped by the

[12] In the analyses that follow, I also explicitly control for other potential determinants of holding periods such as stock and investor characteristics.

[13] This approach follows Seru, Shumway and Stoffman (2008). I obtain similar results by defining the holding period as the time period until all positive positions are closed, as in Feng and Seasholes (2005).

underlying stocks' liquidity.[14] The median holding period is shorter for stocks held in retirement accounts. Investors who are older and who have less wealth invested in the market have shorter holding periods. There is also a strong relationship between holding periods and liquidity of stocks traded by the investors in the dataset.

To explore this relationship further, I rank and assign the 806,404 holding period observations to ten groups based on the length of the holding period. For the stocks in each group, I then calculate averages for the liquidity measures, price, and market capitalization. The liquidity measures are calculated as of the purchase day, by averaging monthly or daily measures over the previous 12 months. The results are reported in Table 2, which show a strong relationship between holding periods and liquidity measures. The relationship is monotonic for most of the measures and is not a simple function of price or market capitalization. The adjusted Amihud illiquidity measure, for instance, increases monotonically from 0.91 to 1.75. There is a 54 basis points (bps) difference in the quoted spread and a 64 bps difference in the realized spread between the highest and the lowest holding period groups.

Figure 3 shows this relationship graphically. I plot Kaplan-Meier survival probabilities for stocks that are in the highest illiquidity decile using the adjusted Amihud illiquidity measure, and for all other stocks in the dataset. The x-axis shows the number of days that have passed since the purchase of a stock, and the two lines plot the probability of an investor holding a stock conditional upon no sale up to that point for the two groups of stocks. Stocks ranked in the highest illiquidity decile have a significantly higher survival probability. The initial univariate results suggest that holding periods are strongly related to measures of baseline transaction costs as predicted in hypothesis *H1a*.

To incorporate stock and investor characteristics, I utilize a hazard model in the analysis of household holding periods.[15] With hazard models, an investor's trade decision can be explicitly modeled by considering the investor's sell-hold decision each

[14] In the figure, a stock is defined as *Illiquid* if it belongs to the lowest liquidity decile of stocks ranked according to the adjusted Amihud illiquidity ratio. *Other* category includes all other stocks not in the lowest liquidity decile.

[15] The hazard framework has been previously used by Seru, Shumway and Stoffman (2008) and Feng and Seasholes (2005) in a similar context to model the disposition effect.

day. In this paper, I use a Cox proportional hazard model with potentially time varying explanatory variables.[16] The hazard model takes the form:

$$\lambda(t) = \lambda_0(t)\exp\left(x(t)'\beta + z'\alpha\right) \tag{6}$$

This is essentially a statistical model that describes how long an investor in the dataset will hold a stock before selling it. The left hand side variable, $\lambda(t)$, is the hazard rate, the probability of selling a stock at day *t* conditional upon holding that stock until that point in time. The explanatory variables are called covariates and can either be static time varying. In equation (6), *x*' represents time-varying covariates and *z*' represents covariates that are fixed over time. $\lambda_0(t)$ 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* (α and β) without specifying a baseline $\lambda_0(t)$ hazard rate.

The static covariates used in this paper are investor and stock characteristics, which are explained in detail in the tables that follow. The only time-varying covariate is a dummy variable that takes on a value of one for each day the stock price trades above its purchase price. This dummy variable measures the disposition effect - a behavioral tendency of investors to sell shares whose price has increased while keeping shares that have dropped in value. Positions that are not closed by the end of the sample period are treated as censored observations. As there is likely to be seasonality in purchases and sales, calendar month dummies are also included as static variables in the hazard regressions.[17] I follow standard reporting conventions and report hazard ratios instead of coefficients from the holding period regressions. The hazard ratio is similar to the odds ratio in binary choice models. It is defined as the ratio of two hazard functions when one of the explanatory variables is changed by one unit holding everything else constant. Since the interpretation of a hazard ratio is more intuitive for dummy variables, I transform the explanatory variables into dummy intervals.

[16] Details about estimating the proportional hazard model can be found in Cox and Oakes (1984).
[17] Open stock positions, for instance, may be closed out in December for tax reasons.

Table 3 shows the results of the hazard regressions. I report results using the adjusted Amihud illiquidity ratio as the transaction costs measure to save space. Similar results are obtained using Pastor and Stambaugh's reversal gamma and the Gibbs estimate of Roll's transaction costs measure. As explained before, the transaction costs measure is calculated by averaging the monthly Amihud illiquidity ratio over the 12 months prior to the purchase date. I rank all stocks by the Amihud illiquidity ratio and create dummy variable (*AdjIlliq Dum*) that takes on a value of one if stock belongs to the highest illiquidity quintile. The hazard ratios corresponding to the dummy variables have an intuitive interpretation. They indicate the probability of a sale (conditional upon no sale up to that point) given that the underlying stock belongs to the highest illiquidity group divided by the probability of a sale given that the stock does not belong to that group. A stock in the highest illiquidity group is 0.6 times as likely to be sold as a stock not belonging to that group.[18] In Model III, I control for investors characteristics and obtain a similar result. As in the univariate analysis, I find that transaction costs are a significant determinant of holding periods of individual investors. The average investor is cognizant of liquidity and pays attention to the transaction costs of the stocks she trades.

The results I report are robust to fixed household effects. One way to capture heterogeneity across households within a hazard framework is to assume a different baseline hazard rate for each household, but compute common coefficients on the explanatory variables. The model is estimated by partial likelihood using the method of stratification. Model II in Table 3 shows that the effect of transaction costs variable increases once I control for fixed household effects. The results suggest that there is variation in holding periods for different stocks for a given household, and that these holding periods are positively related to transaction costs.

I find support for the hypothesis (*H1b*) that the correlation between holding periods and transactions costs increases with investor sophistication and experience. Characteristics we associate with investor sophistication are correlated with shorter holding periods. However, as evidenced by the hazard ratios on the interaction terms (Model IV in Table 3), those who are sophisticated tend to pay attention to the transaction

[18] A stock in the lowest illiquidity group, on the other hand is 1.2 times more likely to be sold than a stock not belonging to that group.

costs of the stocks they trade. Individuals, who are professionals, who have traded options or foreign securities or who have held short positions, have holding periods that are positively correlated with transaction costs. Those who hold mutual funds, on the other hand, have holding periods that are negatively correlated with transaction costs. Individuals who are retired and individuals who trade stocks in their retirement account are more sensitive to transaction costs. In addition, households who have more concentrated portfolios pay more attention to the liquidity of the underlying stocks they trade.

To explore the role of investor sophistication further, I create a numeric variable to proxy for the level of investor sophistication. The *Sophistication* variable starts at a value of zero and is increased by one for each characteristic that one would associate with investor sophistication. I follow Goetzmann and Kumar (2008) and assume that financial sophistication is correlated with education and resources available to an investor. I also use information contained in investors' trades. Table 4 describes the criteria used to construct the *Sophistication* variable. I run the same hazard regression as before (Model I in Table 3), but instead of pooling across all investors, I run a separate regression for each group of investors who have the same *Sophistication* value. For instance, all investors with a *Sophistication* value equal to six would be one group. Figure 4 plots the hazard ratios on the *AdjIlliq Dum* variable for the different groups of investors ranked by S*ophistication*. The relationship between holding periods and transaction costs is stronger for more sophisticated households. The relationship is *negative* for households that are least sophisticated, and there is a monotonic increase in the strength of this relationship as we go from the lowest sophistication group to the highest. In Table 4, I report similar result pooling all investors together. I create a dummy variable (*Sophistication > 3 Dum*) that takes on a value of one if the *Sophistication* value for a given household in the dataset is greater than three. An investor who is sophisticated is 0.4 times as likely to sell an illiquid security at a given point in time, compared to an unsophisticated investor who is 0.6 times as likely to sell an illiquid security.

Although the differences in holding periods for stocks with different liquidity levels are significant, they are substantially lower than the calibrated values in Vayanos (1998) and Constantinides (1986). Vayanos, for instance, predicts an increase in holding period of 6 years when transaction costs increase from 0.5% to 2%. In comparison, a similar increase in transaction costs would increase the holding period of investors by about 190 trading days in the dataset used in this paper. The empirical results are closer to the calibrated values in Lo, Mamaysky and Wang (2006) who predict a similar change in holding periods as in this paper. The results in this section suggest that models that incorporate potentially exogenous liquidity or trading needs are more likely to be representative of actual investor behavior. The results also offer a potential explanation for the discrepancy between the empirically observed liquidity premium and the one predicted by the models in which the holding periods are endogenously determined as in Vayanos (1998) and Constantinides (1986).

4.2. Robustness Checks

To make sure the results are robust to underlying stock characteristics, I include book-to-market, size and momentum characteristics in the hazard regressions. As before, to get a more intuitive interpretation of the results, each year I segment stocks into quintiles based on these stock characteristics. Dummy variables are created and take on a value of one if a stock in the dataset falls into one of the five groups. These characteristics are calculated based on the information available at the beginning of the month in which a sale is made. Table 5 summarizes the results from hazard regressions using these characteristics. The transaction costs measure remains significant after I control for stock characteristics, while the economic and statistical significance of stock characteristics is reduced once I control for liquidity. On average, households tend to hold value and small stocks longer. Relationship between momentum and holding period appears to be U-shaped, but it is more significant at the high return end. A stock belonging to the highest momentum quintile is 1.4 times more likely to be sold conditional on no sale up to that point in time.

The disposition effect (Shefrin and Statman 1985), the tendency of individual investors to hold on to losing stocks too long and to sell winners too quickly, has been shown to be a significant driver of trading behavior in a variety of contexts for both individual and institutional investors. If the disposition effect is the main driver of a decision to buy/sell (Grinblatt and Kellaharjou 2001), then the holding period and the liquidity of a stock would be determined to some extent by how much the stock's current price is above the investors' weighted average purchase price for that stock. Given the robust and significant relationship that has been established in the literature between trading decisions and the disposition effect, and given its close relation to liquidity, I use the disposition effect as a control in the hazard regressions. To do this, as mentioned earlier, I create a time-varying covariate to capture the disposition effect. A dummy variable (*Disp Dum*) is set to one for each day a stock in an investor's portfolio trades above its purchase price. I run the same hazard model as before, but now I include the *disp* variable as a time-varying covariate. The results are provided in Table 5. Using household level controls, I find that an individual is 1.8 times more likely to sell a stock when it is trading above its purchase price than when it is not. The transaction costs variable is significant after controlling for the disposition effect, but is not able to explain away this effect. It is also worth noting that the interaction term is positive, indicating that the disposition effect is stronger among less liquid stocks. Households are more likely to sell an illiquid stock that is trading above the purchase price than one that is not illiquid.

Existence of asymmetric information complicates the analysis. It is not entirely clear how aggregate asymmetric information for a given security would affect its average holding period. On the one hand, one can think of asymmetric information as a component of transaction costs, which investors take into account in selecting which securities to hold. On the other hand, if investors trade for both liquidity and information reasons, allocational inefficiencies could reduce the correlation between holding periods and liquidity (Garleanu and Pedersen 2007). I control for aggregate asymmetric information in a given security by including the probability of information based trading (PIN) measure (Easley et al. 1997) calculated from intra-day data.[19] As before, I compute an annual PIN dummy variable for each stock in the dataset. *PIN Dum* takes on a value of one if the stock is in the highest PIN group. The results appear in Table 5 under Model V. The PIN measure significantly reduces the holding period of investors.

[19] A detailed description is contained in Easley, Hvidkjaer and O'Hara (2004). The data is provided by Soeren Hvidkjaer at http://www.smith.umd.edu/faculty/hvidkjaer/pin1983-2001.zip.

The transaction costs measure, however, does not lose its economic or statistical significance.

As an additional control, I also remove potentially informative trades from the sample. To control for information at the investor level, I run the same model as in the previous section, but remove from the sample trades that may have been conducted for informational reasons. To identify trades that are not motivated by liquidity needs, I follow the same approach in Stoffman (2007). If an individual investor sells his holdings of one security and then immediately uses the proceeds to buy another security, it is unlikely that the particular trade is motivated by liquidity needs. I thus exclude trades that are one trading day apart and for which differences in the values of the trades are less than 5%. Model I in Table 5 shows the results from the hazard regression with these trades removed from the sample. The prior results become stronger when I exclude these potentially informative trades from the dataset.

4.3. Portfolio Level Analyses

I have thus far examined trading decisions of households at the transaction level. I now consider liquidity decisions at the portfolio level. As Amihud and Mendelson (1986) argue, it makes sense for investors with longer holding periods to hold more illiquid stocks in their portfolio since those investors would face lower amortized transaction costs. In this section, I analyze the determinants of overall liquidity of household portfolios and examine how portfolio liquidity is related to households' average holding periods.

Portfolio liquidity is calculated on a monthly basis using position data reported at month end:

$$PIlliq_{i,t} = \frac{\sum_{k=1}^{N} \frac{AdjIlliq_{i,t}^{k}}{MktIlliq_{t}} \times \left| Eq_{i,t}^{k} \right|}{\sum_{k=1}^{N} \left| Eq_{i,t}^{k} \right|} \tag{7}$$

Above, $Eq_{i,t}^{k}$ is the value of stock *k* in household *i*'s portfolio at time *t*, and $AdjIlliq_{i,t}^{k}$ is the adjusted Amihud illiquidity measure of stock *k* in month *t*. $MktIlliq_{t}$ is the market illiquidity, calculated as the equal weighted average *AdjIlliq* of all stocks in month *t*. Since average liquidity varies over time, $MktIlliq_{t}$ is used as an adjustment factor as in Amihud (2002). I average the $PIlliq_{i,t}$ over the sample period to compute an average portfolio illiquidity for each household. Households hold mostly liquid stocks in their portfolio. If we were to rank all stocks by the *AdjIlliq* measure, assign them to percentile ranks, and then calculate a weighted average illiquidity rank for the stocks in an investor's portfolio, 50% of the households would have an average portfolio illiquidity rank that is in the bottom 8th percentile and 75% of the households would have an average portfolio illiquidity rank that is in the bottom 20th percentile.

I calculate a holding period for each household by averaging the holding period for the transactions made by that household. In calculating the average holding periods, I treat positions that are not closed by the end of the sample period as censored. The cross-sectional average and median holding period across households are 437 and 348 trading days respectively.[20] Figure 5 shows the distribution of the average holding periods of households calculated based on transactions that are closed by the end of the sample period, as well as the distribution of holding periods calculated taking into account transactions that are not closed and treated as censored.

Table 6 shows the results from regressing average portfolio liquidity on household holding periods and household characteristics:

$$PIlliq_{i} = \beta_{0} + \beta_{HP} HP_{i} + \sum_{k=1}^{K} \beta_{k} InvCh_{i,k} + \varepsilon_{i} \qquad (8)$$

In equation (8), $PIlliq_{i}$ is the average portfolio illiquidity of household *i*. HP_{i} is the average holding period of household *i*, and $InvCh_{i,k}$ is the k^{th} demographic characteristic of household *i* described in detail in Table 6. Holding period is a statistically significant

[20] The average and median holding period considering only positions that are closed (e.g. ignoring censored observations) are 217 and 168 trading days respectively.

determinant of portfolio liquidity. Given that the median and the 75th percentile adjusted portfolio illiquidity, $PIlliq_i$, across households is 0.037 and 0.105 respectively, what I report is also an economically significant relationship. In Model II, I show that households with higher amounts of wealth invested in the stock market hold more liquid stocks in their portfolio. The same is true for individuals who are older and retired. Investors who hold less diversified portfolios hold more liquid stocks in their portfolios. Overall, the portfolio level results are consistent with the earlier results and hypothesis *H1a*.

5. Holding Periods and Returns

5.1. Amortized Transaction Costs and Returns

In this section, I study the implications of liquidity decisions of individual investors on investment performance. More specifically, I test hypothesis *H2a* outlined in Section 2. The liquidity premium in Amihud and Mendelson (1986) is driven by rents earned by investors who have longer investment horizons. These investors can amortize transaction costs over a longer expected time period and therefore require a lower compensation for holding assets with higher transaction costs. Illiquid assets are shunned by investors who have a shorter time horizon and heavily discounted by them. As a result, long-term investors who bear these costs less frequently earn rents above and beyond the amortized costs of transacting these assets.

I calculate a holding period for each transaction in the dataset that is closed-out by the end of the sample period. I then calculate cumulative raw returns and returns in excess of size, book-to-market and momentum matched portfolios, as in Daniel et al. (1999), over the holding period for each transaction. Characteristics-adjusted excess returns are calculated to make sure that the differences in returns are not driven by differences in stock characteristics.[21] To be able to make comparisons across different holding periods, I calculate average daily returns from cumulative raw and excess returns as:

[21] In the Amihud and Mendelson (1986) model, investors are risk-neutral and in the absence of transaction costs all securities would earn the risk free rate in equilibrium.

$$avgr_i = \sqrt[1/HP]{\prod_{d=1}^{HP} 1 + r_{i,d}} - 1 \tag{9}$$

HP is the holding period measured in days, and $r_{i,d}$ is the daily raw or characteristics-adjusted excess return for transaction *i* in day *d.* I also compute 1, 6, and 12 month raw and excess returns starting from the day of purchase. Transaction costs consist of round trip commissions divided by the value of purchases and sales, as well as the *realized* bid-ask spread for purchases and sales, as described in Section 3. Transaction costs are divided by the holding period to arrive at amortized transaction costs. Consistent with Barber and Odean (2000), I find that on average, each transaction costs one percent in bid-ask spread and 1.4 percent in commissions. In the analyses that follow, I exclude transactions with a holding period of less than two days and stocks priced below two dollars.

I rank all transactions by the holding period and place them into five groups. I then average returns for the transactions in each group. The results are reported in Table 7.[22] In the lowest holding period group, stocks are held on average for 10 days and earn 34.21 basis points (bps) per day before transaction costs. In contrast, stocks in the highest holding period group are held on average for 543 days and earn 2.31 bps per day before transaction costs. Average characteristics-adjusted excess returns are 20.65 bps and -3.59 bps per day before transaction costs, respectively, for the two groups. Thus, short-term traders earn greater daily returns before transaction costs than long-term traders. Short term traders also earn greater 1, 6 and 12 month returns before transaction costs. Once I control for transaction costs, however, the picture changes. For the lowest holding period group, the average return minus amortized commissions and bid-ask spreads is 0.39 bps per day, compared with a net return of 1.14 bps per day for the highest holding period group. Moreover, characteristics-adjusted excess returns are negative for all groups after controlling for transaction costs, but significantly more so for the low holding period group. The difference in returns between the lowest and highest holding period groups is significant. These results are consistent with hypothesis *H2a* outlined in Section 2, in the

[22] Results are reported at the transaction level. I obtain similar results if I aggregate to the household level.

sense that the returns, net of transaction costs, for households with longer holding periods are higher than for households who have shorter holding periods. The relationship for raw returns, however, is not monotonic.

Since I am examining transaction returns as opposed to returns for the whole portfolio, the results could be biased if only profitable trades are closed out producing a disposition effect. In other words, there might be an upward bias for short-term trades, since they may consist mostly of positions that are closed out because the prevailing price is above the purchase price. I consider returns for fixed holding periods from the day of purchase (1, 6, 12 month returns are also reported in Table 7). However, this gets us away from the notion of holding period returns. As a result, I also remove from the sample those households with a strong tendency to close out positions that trade above the purchase price. To identify these households, I split the dataset into two equal time periods and use the first period (from 1991 to 1993) to calculate coefficients on the *disp* variable explained in Section 4. I eliminate households with a positive *disp* coefficient calculated with a 10% confidence level or higher. To make sure that I do not introduce a new bias I also eliminate households with a significant negative *disp* coefficient. I use the second time period (from 1994 to 1996) to calculate holding period returns and amortized transaction costs as described earlier. The results are in Panel B of Table 6. Holding period raw and characteristics-adjusted excess returns are now more uniform. Differences in raw returns between the high and low holding period groups are not significant. There is now a monotonic relationship in returns net of amortized transaction costs across holding period groups, consistent with hypothesis *H2a*.

5.2. Liquidity Decisions and Returns

There are cross-sectional differences in the correlation between holding periods and transaction costs across households. As described in Section 2, this correlation may impact portfolio performance of households on a gross and a net basis. First, households that do not pay attention to transaction costs would be expected to pay higher transaction costs, generating lower *net* returns. Second, a negative correlation between holding periods and transactions costs could also indicate low levels of sophistication and market knowledge, resulting in lower *gross* returns. To identify the two types of households, I use the same hazard model as before, but now instead of pooling across all households, I estimate the coefficient on the transaction costs variable for each household separately. I then use the correlation between holding periods and transaction costs as a proxy for how much each investor pays attention to transaction costs. In order to obtain robust estimates, I require that households make at least 50 round-trip trades over the sample period, and I only keep estimates that are calculated with a 10% confidence level or higher.[23] The summary statistics for the transaction costs coefficient calculated from household level hazard regressions are reported in Table 8. For the majority of households in the dataset (over 60%), the correlation between holding periods and transaction costs is positive. Most investors pay attention to the liquidity level of stocks they trade.

The relationship between holding periods and transaction costs has strong implications for investment performance. I form two groups based on the sign of the coefficient on the transaction costs variable, and calculate 1, 6 and 12 month and holding period returns for each transaction as described in the previous section. I then calculate averages for the two groups. The results are in Table 9. There is a stark difference in the investment performance between the two groups. Households that pay attention to transaction costs earn about 20.5 bps in gross returns and 10.7 bps in characteristics-adjusted excess returns each day, compared to 0.1 bps in gross returns and -6.6 bps in excess returns each day for households that do not. Households that pay attention to transaction costs pay less in amortized spreads and have higher net returns and net characteristics-adjusted excess returns. They earn 7.1 bps per day in net returns, compared to a loss of -10.9 bps per day for households whose holding periods are negatively related to transaction costs. The differences in returns are all statistically significant except for the one month returns. Since the differences are significant for both gross and net returns, the positive relationship between holding periods and transaction costs is consistent with the hypothesis (*H2b)* that investors who pay attention to liquidity earn greater *gross* and *net* returns.

[23] I obtain similar results using 20 or 30 trades instead of 50 trades.

6. Individual Investors and Demand for Liquid Securities

6.1 Common Demand for Liquid Securities

In this section, I extend the analysis to consider how households as a group make liquidity decisions over time. As described in Section 2, commonality in liquidity can arise from investors demanding liquidity at the same time. Increase in uncertainty about changes in future income or wealth, for instance, can cause investors to tilt their portfolios towards more liquid assets at the same time. To test whether there is systematic variation in the trades of liquid assets, I employ a similar methodology used in Kumar and Lee (2006) and Barber, Odean and Zhu (2003), who investigate correlation in the trades of individual investors. Since I make comparisons over time under different regimes of aggregate liquidity, I consider stock liquidity rankings instead of stock liquidity levels. Each month, I rank stocks based on the adjusted Amihud illiquidity measure and assign them to percentile ranks. A stock ranked in the 100th percentile would be the most illiquid stock in a given month. Similarly, a stock ranked in the 1st percentile would be the most liquid.

For groups of non-overlapping investors, *G*, I compute a time series of normalized differences in the liquidity ranks of stocks purchased and sold:

$$IlliqBSI_t^G = \frac{\sum_{i \in G} BV_t^i \times BAdjIlliqRank_{t-1}^i - \sum_{i \in G} SV_t^i \times SAdjIlliqRank_{t-1}^i}{\sum_{i \in G} BV_t^i \times BAdjIlliqRank_{t-1}^i + \sum_{i \in G} SV_t^i \times SAdjIlliqRank_{t-1}^i} \quad (10)$$

where BV_t^i and SV_t^i are the total value of purchases and sales, respectively, for investor *i* in month *t*. $BAdjIlliqRank_{t-1}^i$ and $SAdjIlliqRank_{t-1}^i$ are the weighted average adjusted illiquidity rank of stock holdings of investor *i* belonging to group *G* in month using one month lagged adjusted illiquidity ranks. $IlliqBSI_t^G$ is similar to a buy-sell imbalance index and indicates whether investors belonging to group *G* are net buyers or sellers of liquid securities in a given month. [24] If the trades of liquid securities are independent

[24] There are number of different ways to formulate this analysis. Another approach would be to examine how investors over time tilt their portfolios towards stocks in the top/bottom illiquidity deciles. Kumar

across households, then purchases and sales of liquid stocks by one group of investors will be uncorrelated with that of another group. To test for this independence, I form 5,000 pairs of non-overlapping investor groups containing 500, 1,000 and 5,000 investors. For each $IlliqBSI_t^G$, I then remove the effects of common dependence due to the market factor and common variation in all household trades by running the following regression:

$$IlliqBSI_t^G = \beta_0^G + \beta_{MKT}^G MKT_t + \beta_{BSI}^G BSI_t + \varepsilon_t^G \tag{11}$$

In the equation above, MKT_t is the month *t* market return in excess of the risk free rate, and BSI_t is the buy-sell imbalance for *all* households in a given month *t*, defined as:

$$BSI_t = \frac{\sum_{i \in N} V_{t,Buy}^i - \sum_{i \in N} V_{t,Sell}^i}{\sum_{i \in N} V_{t,Buy}^i - \sum_{i \in N} V_{t,Sell}^i} \tag{12}$$

$V_{t,Buy}^i$ and $V_{t,Sell}^i$ are the total value of purchases and sales, respectively, of investor *i* in month *t*. I aggregate over all *N* investors. The reason for this regression is to remove the common component in the households' net demand for liquid securities due to market movements and changes in overall household demand unrelated to liquidity. I then compute correlations of the residuals, ε_t^G, for different pairs of investor groups.

The results are reported in Table 10. The correlation values range from 18% to 32% depending on the number of investors used in the simulation. All correlations are statistically different from zero. These results suggest the existence of a systematic component in the demand for liquid securities across households. The results support hypothesis *H3a*, that there is systematic variation in households' trades of illiquid securities.

6.2 Aggregate Market Liquidity and Household Demand for Liquid Securities

---

(2009) uses this methodology to investigate dynamic style preferences of individual investors. I obtain similar results examining shifts in portfolio positions as in Kumar (2009).

As mentioned in Section 2, a number of papers treat retail investors as noise traders providing constant liquidity to the market. However, if there is systematic variation in the demand for liquid assets by individual investors, as I have shown in the previous section, then their role as liquidity providers to the rest of the market is not clear. In fact, changes in aggregate liquidity can arise endogenously from correlated trading by individual investors. In this section I investigate how this systematic demand for liquid securities varies with changes in aggregate market liquidity. I test whether there is a *flight to liquidity*, and examine if a subset of individual investors provide liquidity to the market by buying illiquid securities during times of low market liquidity.

I calculate monthly market liquidity as the equal-weighted average of the adjusted Amihud illiquidity ratio for all stocks in a given month (as in Amihud 2002 and Acharya and Pedersen 2005).[25] As before, since I make comparisons over time under different regimes of aggregate liquidity, I consider the liquidity rankings of stocks instead of their liquidity levels. For *all* households, I compute difference in the liquidity ranks of stocks purchased and sold in a given month as:

$$IlliqBSI_t^{ALL} = \frac{\sum_{i \in N} BV_t^i \times BAdjIlliqRank_{t-1}^i - \sum_{i \in N} SV_t^i \times SAdjIlliqRank_{t-1}^i}{\sum_{i \in N} BV_t^i \times BAdjIlliqRank_{t-1}^i + \sum_{i \in N} SV_t^i \times SAdjIlliqRank_{t-1}^i} \quad (13)$$

The variables are defines as in equation (10), but now we compute the sum over all *N* s. Figure 6 plots $IlliqBSI^{ALL}$ and the aggregate market level of illiquidity, *MktIlliq*, over the sample period. In the figure, the period with low market liquidity corresponds with the Mexican peso crises in 1994. The correlation between $IlliqBSI^{ALL}$ and *MktIlliq* is -35%. Individual investors tend to buy liquid stocks and sell illiquid stocks when market liquidity is low.

I split the data into five equal time periods ranked by the aggregate level of market illiquidity. The first time period corresponds to the 34 months with the lowest level of market illiquidity, and the last period to 34 months with the highest level. Table 11

[25] I obtain qualitatively similar results if I use the Pastor and Stambaugh (2003) liquidity measure. The correlation between the measure used in this paper and the Pastor and Stambaugh measure is 30%.

reports the differences in the illiquidity ranks of stocks bought and sold during these five time periods, and also during the month corresponding to the highest level of market illiquidity. When market illiquidity is at its highest point during the 1991 to 1994 period, the difference in the illiquidity rank of the stocks purchased and sold by households is 1.1. When one considers the fact that 50% of the households have an average portfolio illiquidity rank that is in the bottom 8th percentile, the differences I report are both economically and statistically significant. The last column shows the differences in illiquidity ranks of stock purchases and sales adjusted for household portfolio level of liquidity. For this adjustment, I subtract the weighted average illiquidity rank of each household's portfolio from the illiquidity rank of stocks transacted by that household. The magnitude of the differences is lower but still significant and consistent with the earlier result that investors tend to purchase more liquid securities when aggregate liquidity is low.

Table 12 shows the results from regressing illiquidity ranks of stocks purchased or sold in a given month on market illiquidity and investor wealth and income. I estimate the following regression:

$$\begin{aligned} &TransAdjIlliqrank_{k,t} = \beta_0 + \beta_1 Buy_{k,t} + \beta_2 Affluent_i + \beta_3 MktIlliqDum_t \\ &+\beta_4 Buy_{k,t} \times Affluent_i + \beta_5 Buy_{k,t} \times MktIlliqDum_t + \beta_6 MktIlliqDum_t \times Affluent_i \quad (14) \\ &+\beta_7 Buy_{k,t} \times MktIlliqDum_t \times Affluent_i \end{aligned}$$

In equation (14), $TransAdjIlliqrank_{k,t}$ is the lagged adjusted illiquidity rank of the underlying stock for transaction *k* in month *t*.[26] To get a more intuitive interpretation of the regression results, I transform the market illiquidity variable into a dummy variable ($MktIlliqDum_t$) that takes on a value of one for the month in which market illiquidity is at its highest during the sample period. $Buy_{k,t}$ is a dummy variable that takes on a value of one if the transaction *k* in month *t* is a purchase, and $Affluent_i$ is a dummy variable that takes on a value of one if investor *i* is in the highest income bracket (>$100,000) and has

[26] In the regressions, I use lagged (previous month's) illiquidity ranks for stocks transacted in a given month. I obtain similar results using contemporaneous illiquidity ranks.

invested more than $100,000 in the stock market during the sample period.[27] Model I in Table 11 shows that on average, when market illiquidity is high, households trade more liquid stocks. The coefficient on the interaction term, $MktIlliq\ Dum \times Buy$, in Model II is negative. Since I am using dummy variables, the coefficient on the interaction term shows how much the illiquidity rank of the stocks purchased are higher or lower than stocks sold during times of low market liquidity. The -1.6 coefficient on the interaction term is economically and statistically significant. Controlling for fixed household effects in Model III slightly reduces the effect to -1.0.

In hypothesis *H3c*, I predict that households with higher levels of wealth and income buy illiquid assets that have dropped in price providing liquidity to the market. The interaction term, $MktIlliq\ Dum \times Buy \times Affluent$, in Model IV in Table 12 is positive. Households with higher incomes and higher amounts invested in the stock market tend to buy more illiquid stocks during times of low market liquidity. The net effect of an increase in illiquidity rank of purchases by *Affluent* households during times of high market illiquidity is 0.93. As before, this result is both economically and statistically significant. The results are consistent with the hypothesis that investors with deep pockets provide liquidity to the market by purchasing illiquid stocks when market liquidity is low.

7. Conclusion

This paper investigates both portfolio and stock level liquidity decisions of 66,000 households from a large discount brokerage. It provides an empirical link between investor decisions and the liquidity premium observed in the market. Three main conclusions follow from the analysis. First, transaction costs are an important determinant of investment policies and trading decisions. 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 various controls, including household and stock characteristics as well as the disposition effect and the level of asymmetric information. The results also hold at the portfolio level.

[27] I obtain similar results if I use a $75,000 or $150,000 cut-off for income and wealth invested in the stock market.

Consistent with the notion of *liquidity clienteles*, investors with longer investment horizons tend to hold more illiquid securities. There is cross-sectional variation in the relationship between holding periods and transaction costs across households, and I find that this relationship is stronger among more sophisticated investors. Second, I show that liquidity decisions have important implications for investment performance. As postulated by Amihud and Mendelson (1986), households with longer holding periods earn significantly higher returns after amortized transaction costs. In addition, households that have holding periods that are negatively related to transaction costs earn, on average, lower *gross* and *net* returns. Finally, this paper shows that there is systematic variation in demand for liquid assets across investors. Consistent with the notion of *flight to liquidity*, households are net demanders of liquid securities during times of low aggregate market liquidity. Households with higher incomes and higher wealth invested in the stock market supply liquidity when market liquidity is low.

References

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

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.

Atkins, A. B. and E. A. Dyl, 1997, "Transactions costs and holding periods for common stocks,"*Journal of Finance,* 52, 309–325.

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*," Quarterly Journal of Economics*, 116, 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., T. Odean, and N. Zhu, 2003, "Systematic noise," Working paper, Haas School of Business, University of California at Berkeley.

Barber, B. M., T. Odean, and N. Zhu, 2006, "Do noise traders move markets?," Working Paper, UC Davis.

Barber, B. M., T. Odean, Y. Lee, and Y. Liu, 2008, "Just how much do investor lose from trade?," Working paper, UC Davis.

Campbell, J. Y., S. J. Grossman, and J.Wang, 1993, "Trading volume and serial correlation in stock returns," *Quarterly Journal of Economics*, 108, 905–939.

Campbell, J. Y., T. Ramadorai, and A. Schwartz, 2007, "Caught on tape: Institutional trading, stock returns, and earnings announcements," Working paper, Harvard University.

Chordia, T., R. Roll, and A. Subrahmanyam, 2000, "Commonality in liquidity," *Journal of Financial Economics,* 56, 309–325.

Chordia, T., A. Subrahmanyam, and V. R. Anshuman, 2001, "Trading activity and expected stock returns," *Journal of Financial Economics,* 59, 3–32.

Chordia, T., S.Huh, and A. Subrahmanyam, 2007, “Theory-based illiquidity and asset pricing,” Working paper, UC Los Angeles.

Comerton-Forde, C., T. Hendershott, C.M. Jones, P.C. Moulton, and M.S. Seasholes, 2008, “Time variation in liquidity: The role of market maker inventories and revenues,” Working Paper, Haas School of Business, University of California at Berkeley.

Constantinides, G. M., 1986, “Capital market equilibrium with transaction costs,” *Journal of Political Economy*, 94, 842–862.

Coughenour, J.F., and M.M. Saad, 2004, “Common market makers and commonality in liquidity," *Journal of Financial Economics*, 73, 37-70.

Coval, J. D., D. A. Hirshleifer, and T. Shumway, 2005, “Can individual investors beat the market?,” Working Paper, University of Michigan.

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

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

Datar, V. T., N. Y. Naik, and R. Radcliffe, 1998, “Liquidity and stock returns: An alternative test,” *Journal of Financial Markets*, 1, 205–219.

Easley, D., N. M. Kiefer, and M. O”Hara, 1997, “One day in the life of a very common stock,” *Review of Financial Studies*, 10, 805–835.

Easley, David, Soeren Hvidkjaer, and Maureen O”Hara, 2002, “Is information risk a determinant of asset returns?,” *Journal of Finance*, 57, 2185–2221.

Eckbo, E., and O. Norli, 2002, “Pervasive liquidity risk”, Working paper, Tuck School of Business.

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.

Garleanu, N. and L. H. Pedersen, 2004, “Adverse selection and the required return,” *Review of Financial Studies*, 17, 643–665.

Goetzmann, W. N., A. Kumar, 2008, “Equity portfolio diversification,” *Review of Finance*, forthcoming.

Griffin, J. M., J. H. Harris, and S. Topaloglu, 2003, “The dynamics of institutional and individual trading,” *Journal of Finance*, 58, 2285–2320.

Grinblatt, M., and M. Keloharju, 2001, “What makes investors trade?,” *Journal of Finance*, 56, 589–616.

Hasbrouck, J., 2005, “Inferring trading costs from daily data: US equities for 1962 to 2001,” Working Paper, New York University.

Hasbrouck, J. and D. Seppi, 2001, “Common factors in prices, order flows, and liquidity,” *Journal of Financial Economics,* 59, 383–411.

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

Huberman, G. and D. Halka, 2001, “Systematic liquidity,” *Journal of Financial Research*, 24, 161–178.

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

Hvidkjaer, S., 2008, “Small trades and the cross-section of stock returns**,”** *Review of Financial Studies*, 21, 1123-1151.

Ivkovi´c, Z., C. Sialm, and S. J. Weisbenner, 2006, “Portfolio concentration and the performance of individual investors,” NBER Working Paper No. W10675.

Ivkovi´c, Z., and S. J. Weisbenner, 2005, “Local does as local is: Information content of the geography of individual investors common stock investments,” *Journal of Finance*, 60, 267–306.

Kaniel, R., G. Saar, and S. Titman, 2006, “individual investor trading and stock returns,” *Journal of Finance*, forthcoming.

Korajczyk, R. A., and R. Sadka, 2008. “Pricing the commonality across alternative measures of liquidity,” *Journal of Financial Economics*, 87, 45-72.

Kumar, A. and M.C. Lee, 2006. "Retail investor sentiment and return comovements," *Journal of Finance*, 61, 2451-2486.

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, 2004, “Explaining the magnitude of liquidity premia: The roles of return predictability, wealth shocks and state dependent transaction costs,” Working Paper, New York University.

Naes, R., and B. A. Odegaard, 2008, “liquidity and asset pricing: Evidence on the role of investor holding period”, Working Paper, Norwegian School of Management.

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

Pastor, L., and R. F. Stambaugh, 2003, “Liquidity risk and expected stock returns,” *Journal of Political Economy*, 111, 642–685.

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

Seru, A., T. Shumway, and N. Stoffman, 2008, “Learning by trading,” Working paper, University of Michigan.

Shefrin, H., and M. Statman, 1985, “the disposition to sell winners too early and ride losers too long: Theory and evidence,” *Journal of Finance*, 40, 777–790.

Stiglitz, J.E., 1989, “Using tax policy to curb speculative short-term trading,” *Journal of Financial Services Research*, 3, 101-115.

Stoffman, N., 2007, “When are individual investors informed?,” Working paper, Ross School of Business, University of Michigan

Stoffman, N., 2008, “Who trades with whom? individuals, institutions, and returns,” Working paper, Ross School of Business, University of Michigan.

Summers, L.H., V.P. Summers, 1989, “When financial markets work too well: A cautious case for a securities transactions tax,” *Journal of Financial Services Research*, 3, 261-286.

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, 1997, “A preferred-habitat model of the term structure of interest rates,” Working paper, London School of Economics.

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

Vayanos, D., 2004, “Flight to quality, flight to liquidity, and the pricing of risk,” Working paper, London School of Economics.

Table 1: Liquidity Measures Summary Statistics

This table reports summary statistics and correlations for the liquidity measures used in this paper. Only stocks that are traded by households in the dataset are considered. Summary statistics and correlations are calculated by pooling annual observations over the 1991-1996 time period. All liquidity measures are annual averages and are defined in the text. Price is the average of the daily closing prices. Mkt Cap is the average market capitalization. PS Gamma is Pastor and Stambaugh (2003) reversal gamma. AdjIlliq is the adjusted Amihud illiquidity ratio. Rolls C is the Bayesian estimate of the Roll (1984) transaction cost measure. The quoted spread is the ratio of the quoted bid-ask spread to the prevailing transaction price. The effective spread is the absolute value of the difference between the transaction price and the quote midpoint, divided by the bid-ask midpoint. The quoted depth is the average of quoted bid-ask lots multiplied by bid-ask quotes. Realized spread is the realized bid-ask spread paid by the investors for each transaction in the dataset. Spreads are reported in basis points.

	PS Gamma	**Roll's C**	**Price $**	**Mkt Cap 000s**	**Amihud AdjIlliq**	**Quoted Spread**	**Effective Spread**	**Quoted Depth**
Mean	26.81	1.73	18.61	1,045	7.23	324.32	205.53	685.72
Median	0.49	1.06	12.50	105	1.66	162.13	95.54	385.06
Std	316.79	2.05	124.04	4,549	9.70	532.61	385.57	898.28
P25	-0.34	0.45	5.00	30	0.38	100.54	54.56	161.69
P75	9.18	2.26	24.50	452	11.64	301.51	186.85	845.93
				Pearson Correlations				
PS Gamma	1.00							
Rolls C	0.15	1.00						
Price	-0.01	-0.07	1.00					
Mkt Cap	-0.02	-0.15	0.09	1.00				
AdjIlliq	0.16	0.76	-0.07	-0.16	1.00			
Quoted Spread	0.17	0.58	-0.30	-0.13	0.54	1.00		
Effective Spread	0.18	0.60	-0.29	-0.12	0.53	0.94	1.00	
Quoted Depth	-0.04	-0.24	0.40	0.61	-0.28	-0.27	-0.25	1.00
				Spearman Correlations				
PS Gamma	1.00							
Rolls C	0.41	1.00						
Price	-0.34	-0.81	1.00					
Mkt Cap	-0.38	-0.83	0.85	1.00				
AdjIlliq	0.39	0.85	-0.76	-0.91	1.00			
Quoted Spread	0.15	0.71	-0.80	-0.76	0.75	1.00		
Effective Spread	0.16	0.72	-0.81	-0.79	0.78	0.97	1.00	
Quoted Depth	-0.13	-0.53	0.61	0.78	-0.77	-0.78	-0.78	1.00

Table 2: Univariate Results

This table presents the univariate results. Transactions in the dataset are ranked by holding-periods and placed into ten groups. Averages for the various liquidity measures for the underlying securities are then calculated for each group. All liquidity measures are annual averages and are defined in the text. Price is the average of the daily closing prices. Mkt Cap is the average market capitalization. PS Gamma is Pastor and Stambaugh (2003) reversal gamma. AdjIlliq is the adjusted Amihud illiquidity ratio. Rolls C is the Bayesian estimate of the Roll (1984) transaction cost measure. The quoted spread is the ratio of the quoted bid-ask spread to the prevailing transaction price. The effective spread is the absolute value of the difference between the transaction price and the quote midpoint, divided by the bid-ask midpoint. The quoted depth is the average of quoted bid-ask lots multiplied by bid-ask quotes. Realized spread is the realized bid-ask spread paid by the investors for each transaction in the dataset. Spreads are given basis points. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	Holding Period	**Price $**	**Mkt Cap 000s**	**PS Gamma**	**Amihud AdjIlliq**	**Roll's C**	**Quoted Spread**	**Effective Spread**	**Quoted Depth**	**Realized Spread**
Low	6	32.89	7,940	1.2743	0.9142	0.6608	118.84	83.61	3,140	62.99
2	20	32.04	7,602	1.7055	0.9943	0.6834	124.19	86.96	3,027	78.27
3	44	31.28	7,833	2.2359	1.0893	0.7054	123.82	86.83	3,058	104.96
4	79	31.52	9,029	2.3783	1.1265	0.7072	119.13	83.38	3,185	112.82
5	127	33.96	9,199	2.7837	1.2606	0.7337	119.24	84.20	3,266	132.60
6	194	31.21	10,513	3.1087	1.2421	0.7312	117.93	83.26	3,415	130.82
7	294	30.36	9,886	3.9784	1.3535	0.7382	115.72	81.57	3,341	139.63
8	470	29.47	10,266	3.6685	1.3819	0.7312	113.94	80.57	3,519	136.90
9	771	31.74	11,434	4.4004	1.4889	0.7425	115.51	81.71	3,748	129.28
High	1225	40.76	11,270	6.4303	1.7578	0.8182	172.55	121.30	2,977	127.18
High – Low	1219***	7.87***	3330***	5.156***	0.8436***	0.1575***	53.71***	37.69***	-162***	64.19***

Table 3: Hazard Regressions

This table reports hazard ratios from the holding period regressions where the conditional probability of sale is the dependent variable. Independent variables consist of a transactions costs measure and a set of investor demographic and trade variables. AdjIlliq Dum is a dummy variable that takes on a value equal to one if a stock in the dataset is in the highest quintile ranked by the adjusted Amihud illiquidity ratio calculated over the previous 12 months prior to a transaction. Age [40-64] Dum is a dummy variable set equal to one if the age of the head of the household is between 40 and 64. Age 65+ Dum is a dummy variable set equal to one if the age of the head of the household is over 64. Income > 75K Dum is a dummy that is set to one if the total annual household income exceeds $75K. Married Dum is a dummy variable set to one if the head of the household is married. Male Dum is set to one if the head of the household is male. Professional Dum and Retired Dum are dummy variables that reflect investors' occupation. Professional Dum is set to one for investors who hold technical and managerial positions and Retired Dum is set to one for investors who are retired. Retirement Account Dum is set to one if the underlying account is a retirement (IRA or Keogh) account. Trade variables are derived from the trades made by investors in the dataset. Short User Dum is set to one if an investor executed at least one short-sell during the sample period. Option User Dum is set to one if an investor has traded in options. Mutual Fund user Dum is set to one if an investor has held mutual funds during the sample period. Foreign User Dum is set to one if an investor made at least one trade in a foreign asset including ADRs, foreign stocks or foreign mutual funds during the sample period. Total Equity > 45K is dummy variable set to one if household's total value of equity invested in the stock market exceeds $45K. Diversification is defined as in Goetzmann and Kumar (2008), and is equal to the sum of the squared value weight of each stock in a household's portfolio. Diversification < 0.3 Dum is dummy variable if this diversification measure for a given household is less than 0.3. Calendar month dummies (not reported) are twelve dummy variables that take on a value of one if the month of the transaction is equal to the month dummy. Robust standard errors are calculated as in Lin and Wei (1989). Ties are handled using the Efron procedure. Wald test is for each additional set of regressors. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	Model I		**Model II**		**Model III**		**Model IV**	
	Haz Ratio	**p-val**	**Haz Ratio**	**p-val**	**Haz Ratio**	**p-val**	**Haz Ratio**	**p-val**
AdjIlliq Dum	0.617***	*<.0001*	0.602***	*<.0001*	0.632***	*<.0001*	0.804**	*0.0470*
Demographic Variables								
Age [40 - 64] Dum					0.988	*0.1698*	0.986	*0.1168*
Age 65 + Dum					0.83***	*<.0001*	0.829***	*<.0001*
Income > 75K Dum					0.921***	*<.0001*	0.921***	*<.0001*
Married Dum					0.945***	*<.0001*	0.945***	*<.0001*
Male Dum					1.101***	*<.0001*	1.101***	*<.0001*
Professional Dum					1.009	*0.3158*	1.01	*0.2319*
Retirement Acct Dum					0.852***	*<.0001*	0.852***	*<.0001*
Retired Dum					1.091***	*<.0001*	1.093***	*<.0001*
Trade Variables								
Foreign securities Dum					1.146***	*<.0001*	1.147***	*<.0001*
Mutual fund user Dum					0.988**	*0.0701*	0.985**	*0.0232*
Option user Dum					1.492***	*<.0001*	1.497***	*<.0001*
Short user Dum					1.968***	*<.0001*	1.976***	*<.0001*
Total Equity > 45K Dum					1.314***	*<.0001*	1.318***	*<.0001*
Diversification < 0.3 Dum					0.705***	*<.0001*		
Interactions								
AdjIlliq Dum * Age [40 - 64] Dum							1.151*	*0.0832*
AdjIlliq Dum * Age 65+ Dum							1.189	*0.1286*
AdjIlliq Dum * Income > 75K Dum							0.933***	*<.0001*
AdjIlliq Dum * Married Dum							1.019	*0.7706*
AdjIlliq Dum * Male Dum							0.975	*0.8240*

AdjIlliq Dum * Professional Dum				0.863**	*0.0364*
AdjIlliq Dum * Retirement Acct Dum				0.959*	*0.0521*
AdjIlliq Dum * Retired Dum				0.858**	*0.0129*
AdjIlliq Dum * Foreign Dum				0.919**	*0.0179*
AdjIlliq Dum * Mutual fund Dum				1.274***	*<.0001*
AdjIlliq Dum * Option user Dum				0.794***	*0.0013*
AdjIlliq Dum * Short user Dum				0.781***	*<.0001*
AdjIlliq Dum * Total Equity				0.854**	*0.5464*
AdjIlliq Dum * Diversification < 0.3 Dum				1.197***	*0.0032*
Household effects	No	Yes	No		No
Calendar month dummies	Yes	Yes	Yes		Yes
Wald test	<.0001	<.0001	<.0001		<.0001

Table 4: Household Sophistication Measure

The top panel lists the criteria used to construct the *Sophistication* variable. This variable is increased by a value of one if an investor in the dataset meets anyone of the criteria listed n the table. The bottom panel reports hazard ratios from the holding period regression, where the conditional probability of sale is the dependent variable. AdjIlliq Dum is a dummy variable that takes on a value equal to one if a stock in the dataset is in the highest quintile ranked by the adjusted Amihud illiquidity ratio calculated over the previous 12 months prior to a transaction. Sophistication > 3 Dum is dummy variable set to one if the Sophistication variable for an investor in the dataset is greater than three. Calendar month dummies (not reported) are twelve dummy variables that take on a value of one if the month of the transaction is equal to the month dummy. Robust standard errors are calculated 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.

Criteria	**Sophistication**
Income > $75K	+ 1
Equity Investments > $45K	+ 1
Investor is a professional	+ 1
Trades Options	+ 1
Trades Foreign Securities	+ 1
Does not invest in Mutual Funds	+ 1
Has held a Short position	+ 1
Portfolio Diversification < 0.3	+ 1

	Haz Ratio	**p-val**
AdjIlliq Dum	0.625***	<*.0001*
Sophistication > 3 Dum	1.110***	<*.0001*
Sophistication > 3 * AdjIlliq Dum	0.714***	<*.0001*
Calendar Month Dummies	Yes	

Table 5: Robustness Checks

This table reports the result of hazard regressions where the holding period is the dependent variable. The independent variables are the transaction costs measure, stock characteristics, the disposition effect proxy, and the PIN measure. AdjIlliq is the average adjusted Amihud illiquidity ratio over a year. Size is the market capitalization. Book-to-market is the book value from the previous fiscal year divided by the current market capitalization. Momentum is the previous 12 month return. PIN is the annual average of probability of informed trading (Easley et al. 1997) variable. Dummy variables (Dum) are created for the transaction costs measure, stock characteristics, and the PIN measure and set to one if a stock is in the highest quintile ranked according to one these variables. For the transaction costs and the PIN measures, stocks are ranked and sorted into quintiles at the beginning of the month of a purchase. The same procedure is repeated for the stock characteristics, but the ranking is done at the beginning of the month when there is a sale. Disp Dum is the disposition proxy. It is a time-varying dummy variable that takes on a value of one if the stock at given day is trading above its purchase price. The investor characteristics are the household demographic and trade variables defined in Table 3. Calendar month dummies (not reported) are twelve dummy variables that take on a value of one if the month of the transaction is equal to the month dummy. Robust standard errors are calculated 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.

	Model I	**Model II**	**Model III**	**Model IV**	**Model V**	**Model VI**	**Model VII**	**Model VIII**	**Model IX**
AdjIlliq Dum	0.587***						0.660***	0.687***	0.695***
	<.0001						*<.0001*	*<.0001*	*<.0001*
Book-to-Market Dum		0.837***					0.913**		
		<.0001					*<.0001*		
Size Sum			1.174***				1.146***		
			<.0001				*<.0001*		
Momentum Dum				1.438***			1.417***		
				<.0001			*<.0001*		
PIN Dum					1.182***			1.229***	
					<.0001			*0.1015*	
Disp Dum						1.810***			1.793***
						<.0001			*<.0001*
AdjIlliq Dum * Disp Dum									1.141***
									<.0001
Calendar month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Investor Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: Portfolio Liquidity and Holding Periods

This table reports the results of regressions using portfolio illiquidity as the dependent variable. The independent variables are investor holding periods and investor characteristics. PIlliq is the average household portfolio illiquidity as defined in Section 4.3. Holding period is the average household holding period. It is calculated by averaging holding periods for all transactions of a given investor. Positions that are not closed-out by the end of the sample period are treated as censored observations. A censored average is calculated assuming a Weibull distribution for the holding period. Investor characteristics are described in Table 3. Robust standard errors are reported below coefficient estimates. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	Model I	**Model II**
Holding Period (years)	0.0515***	0.0631***
	0.0079	*0.0152*
Age		-0.0012***
		0.0002
Income		0.0002
		0.0008
Married Dum		-0.0219
		0.0007
Professional Dum		-0.0205***
		0.0069
Retired Dum		-0.0181**
		0.0099
Male Dum		0.0591***
		0.0097
Foreign securities Dum		0.0487***
		0.0079
Mutual fund user Dum		0.001
		0.0057
Option user Dum		0.0709***
		0.0096
Short user Dum		0.0122***
		0.0065
Log Total Equity		-0.0981***
		0.0024
Diversification		-0.0334***
		0.0113
N	63,024	19,746
Adj R^2	0.01	0.09

Table 7: Holding Period Returns

This table reports transaction returns to holding period groups. Holding period is defined as the time period from the first purchase to the first sale of a security. Transactions are ranked and put into holding period quintiles. 1, 6, and 12 month returns are calculated starting from the date of purchase. Holding period returns are average daily returns (reported in basis points) over the holding period. Excess returns are returns net of characteristics matched portfolios, as in Daniel et al. (1997). Amortized spread is the realized spread (as defined in Table 2) divided by the holding period. Amortized commission is the round-trip commission divided by the holding period. Transactions with a purchase or sale price less than $2, and holding periods less than 2 days, are excluded from the sample. Panel B reports returns for a sub-sample of the households in the 1994-1996 time period. The 1991-1993 time period is used to calculate a coefficient on the *disp* variable for each household in the dataset. Households with a positive *disp* coefficient significant at the 10% level are removed from the sample. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: Returns to Holding Period Groups

	Low	**2**	**3**	**4**	**High**	**High - Low**
1 Month Ret	0.045	0.036	0.011	0.004	0.001	-0.044***
1 Month Excess Ret	0.018	0.010	-0.006	-0.012	-0.013	-0.031***
6 Month Ret	0.079	0.112	0.132	0.054	0.008	-0.071***
6 Month Excess Ret	-0.009	0.011	0.025	-0.031	-0.055	-0.045***
12 Month Ret	0.148	0.187	0.200	0.188	0.056	-0.092***
12 Month Excess Ret	-0.014	0.007	0.012	-0.003	-0.081	-0.067***
Holding Period Ret (bps)	34.211	15.080	8.085	4.116	2.307	-31.904***
Holding Period Excess Ret (bps)	20.648	4.446	0.045	-2.778	-3.587	-24.235***
Holding Period Net Ret (bps)	0.386	3.280	2.603	1.358	1.137	0.751*
Holding Period Net Excess Ret (bps)	-13.177	-7.354	-5.436	-5.537	-4.757	8.420***
Amortized Spread (bps)	5.257	3.063	1.501	0.721	0.264	-4.993***
Amortized Commission (bps)	28.568	8.737	3.981	2.037	0.906	-27.662***
Holding Period	10	36	87	192	543	533***

Panel B: Bias Adjusted Returns to Holding Period Groups

	Low	**2**	**3**	**4**	**High**	**High - Low**
1 Month Ret	0.016	0.027	0.020	0.010	0.002	-0.014***
1 Month Excess Ret	-0.002	0.006	0.001	-0.006	-0.009	-0.007***
6 Month Ret	0.049	0.078	0.109	0.119	0.051	0.002
6 Month Excess Ret	-0.034	-0.013	0.007	0.011	-0.032	0.002
12 Month Ret	0.112	0.153	0.201	0.232	0.187	0.075***
12 Month Excess Ret	-0.038	-0.014	0.004	0.013	-0.024	0.014***
Holding Period Ret (bps)	1.383	2.626	4.739	5.031	4.371	2.988
Holding Period Excess Ret (bps)	-2.402	-4.392	-2.846	-2.547	-3.517	-1.115
Holding Period Net Ret (bps)	-38.105	-12.659	-2.171	1.514	2.676	40.781***
Holding Period Net Excess Ret (bps)	-41.889	-19.677	-9.756	-6.065	-5.212	36.677***
Amortized Spread (bps)	5.588	3.844	1.819	0.886	0.377	-5.210***
Amortized Commission (bps)	33.900	11.441	5.091	2.631	1.318	-32.582***
Holding Period	7	24	59	125	309	302***

Table 8: Household Transaction Costs Coefficient Estimates

This table reports summary statistics of the transaction costs coefficient, which is calculated from household level hazard regressions described in Section 5.2. *AdjIlliq* variable is used as the transaction costs measure. To get robust estimates, households are required to have made at least 50 trades during the sample period to be included in the analysis. The summary statistics for the coefficients calculated with at least 10% statistical significance are reported in the second column.

	All Obs	**Obs Significant at >10%**
Mean	-0.3002	-0.5834
Median	-0.1089	-0.2752
Std Dev	4.8435	7.5727
Skew	-29.745	-20.165
Kurtosis	1170.52	507.27
P5	-1.1015	-1.5748
P25	-0.3366	-0.5266
P75	0.1188	0.3018
P95	0.6860	1.2017

Table 9: Transaction Costs and Holding Period Returns

This table reports transaction returns to two groups formed based on the sign of the transaction costs coefficient, which is calculated from household level hazard regressions described in Section 5.2. *AdjIlliq* variable is used as the transaction costs measure. To get robust estimates, households are required to have made at least 50 trades during the sample period to be included in the analysis. 1, 6, and 12 month returns are calculated starting from the date of purchase. Holding period returns are average daily returns (reported in basis points) calculated from the first purchase of a security to the first sale. Excess returns are returns net of characteristics matched portfolios, as in Daniel et al. (1997). Amortized spread is the realized spread (as defined in Table 2) divided by the holding period. Amortized commission is the round-trip commission divided by the holding period. Transactions with a purchase or sale price less than $2, and holding periods less than 2 days, are excluded from the sample. Panel B reports returns for the full sample, and Panel A reports returns where the coefficient on the AdjIlliq variable is calculated with at least 10% significance. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: Observations with AdjIlliq Coefficient at >10% Significance			
	Positive	**Negative**	**Positive - Negative**
1 Month Ret	0.018	0.018	0.001
1 Month Excess Ret	-0.001	-0.001	0.001
6 Month Ret	0.079	0.066	0.013***
6 Month Excess Ret	-0.010	-0.020	0.01***
12 Month Ret	0.161	0.132	0.029***
12 Month Excess Ret	-0.010	-0.035	0.025***
Holding Period Ret (bps)	20.450	0.122	20.327***
Holding Period Excess Ret (bps)	10.756	-6.564	17.32***
Holding Period Net Ret (bps)	7.077	-10.950	18.027***
Holding Period Net Excess Ret (bps)	-2.617	-17.636	15.019***
Amortized Spread (bps)	0.675	2.202	-1.527***
Amortized Commission (bps)	12.697	8.870	3.827***
Holding Period	100	157	-57***
Panel B: All Observations			
	Positive	**Negative**	**Positive - Negative**
1 Month Ret	0.018	0.017	0.001**
1 Month Excess Ret	-0.001	-0.002	0.002**
6 Month Ret	0.079	0.070	0.009***
6 Month Excess Ret	-0.010	-0.019	0.009***
12 Month Ret	0.162	0.146	0.016***
12 Month Excess Ret	-0.009	-0.027	0.018***
Holding Period Ret (bps)	16.909	4.125	12.785***
Holding Period Excess Ret (bps)	7.621	-3.542	11.163***
Holding Period Net Ret (bps)	4.228	-7.570	11.798***
Holding Period Net Excess Ret (bps)	-5.060	-15.236	10.176***
Amortized Spread (bps)	0.942	2.259	-1.317***
Amortized Commission (bps)	11.739	9.435	2.304***
Holding Period	116	147	-32***

Table 10: Common Demand for Liquidity

This table reports correlation statistics from three different simulations that test for a systematic component in the demand for liquid assets across households. A pair of non-overlapping investor groups containing *N* investors (where *N* = 500, 1,000 and 5,000) is selected from the dataset. The normalized difference in the liquidity ranks of stocks the investors in each group purchase and sell each month are calculated (*IlliqBSI* variable in Equation 10). IlliqBSI for each investor group is regressed on the market factor and the aggregate buy-sell imbalance to remove the common variation in all household trades unrelated to liquidity. A time series correlation of the residual from the regression is calculated between two groups of investors. The same procedure is repeated 5,000 times. The summary statistics for the 5,000 simulated correlations are reported below.

# of Investors	Mean	Median	Std Dev	t-value
500	0.1782	0.1559	0.3005	41.95
1000	0.2108	0.2409	0.2790	53.43
5000	0.3799	0.3826	0.1636	164.18

Table 11: Illiquidity Rank of Transactions

This table reports the differences in the adjusted illiquidity ranks of household purchases and sales of securities under different levels of aggregate market illiquidity. Market illiquidity is calculated as the equal-weighted average of the adjusted Amihud illiquidity ratio of all stocks in a given month. The sample period is broken into five equal time periods determined by the level of market illiquidity, ranked from 'Low' to 'High' in the table. 'MAX' is the month corresponding to the highest level of market illiquidity. Stocks are ranked each month based on the adjusted Amihud Illiquidity measure and assigned to percentile ranks. The adjusted illiquidity rank of purchases and sales and the difference between purchases and sales are reported for five different levels of aggregate liquidity and for the month in which the market illiquidity is at its highest. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Market Illiquidity	Buy/Sell	N Obs	Adj Illiquidity Rank	HH demeaned Adj Illiquidity Rank
Low	Buy	188,601	16.71	0.94
	Sell	155,111	16.05	0.24
	Diff		0.66***	0.7***
2	Buy	226,817	15.87	0.29
	Sell	185,471	15.86	-0.03
	Diff		0.01	0.32***
3	Buy	186,929	16.00	0.43
	Sell	155,989	15.44	-0.18
	Diff		0.56***	0.61***
4	Buy	244,573	15.97	0.36
	Sell	201,018	15.44	-0.31
	Diff		0.53***	0.67***
High	Buy	215,823	16.35	0.58
	Sell	174,064	17.21	0.99
	Diff		-0.86***	-0.41***
MAX	Buy	11,436	14.94	-0.20
	Sell	7,659	16.06	0.27
	Diff		-1.13***	-0.47*

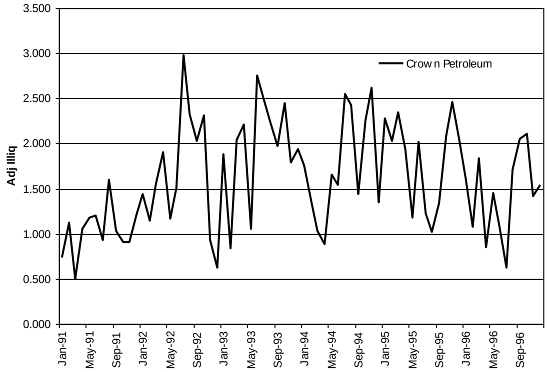
Table 12: Market Liquidity and Liquidity of Transactions

This table reports the result of regressions using the illiquidity rank of the security that is purchased or sold as the dependent variable. The independent variables are aggregate market illiquidity and investor income and wealth. Market illiquidity is calculated as the equal-weighted average of the adjusted Amihud illiquidity ratio of all stocks in a given month. MktIlliq is a dummy variable that takes on a value of one if the aggregate market illiquidity is in the lowest month during the sample time period. Buy is a dummy variable that takes on a value of one if the transaction is a purchase. Affluent is a dummy variable that takes on a value of one if the investor is in the highest income bracket (>$100,000) and has invested more than $100,000 in the stock market during the sample period. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	Model I	**Model II**	**Model III**	**Model IV**
MktIlliq	-0.8688 ***	0.0715	0.039	0.5357
	0.1509	0.2380	0.2108	0.3710
Buy		0.2892***	0.2961***	0.3174***
		0.0301	0.0267	0.0433
Buy * MktIlliq		-1.5957***	-1.009***	-2.6296***
		0.3078	0.2710	0.4817
Buy * MktIlliq * Affluent				2.1666***
				0.8313
Affluent				-1.2371***
				0.0210
Buy * Affluent				-0.7172
				0.6384
Affluent * MktIlliq				-0.2302***
				0.0782
Adj R^2	0.01	0.01	0.01	0.08
Household Effects	No	No	Yes	No

Figure 1: Illiquidity Ratio

This figure shows the adjusted Amihud illiquidity ratio for IBM and Crown Petroleum Corp from Jan. 1991 to Dec. 1996.



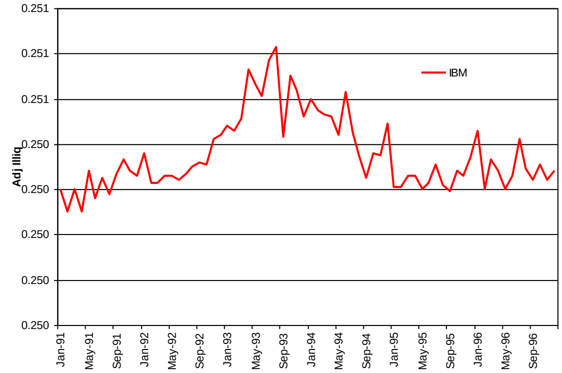


Figure 2: Holding Periods of Households

This figure shows the median holding period for various investor and stock groups. 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 lowest liquidity decile of stocks ranked according to the adjusted Amihud illiquidity ratio. The holding period is calculated only for positions that are closed-out by the end of the sample period.

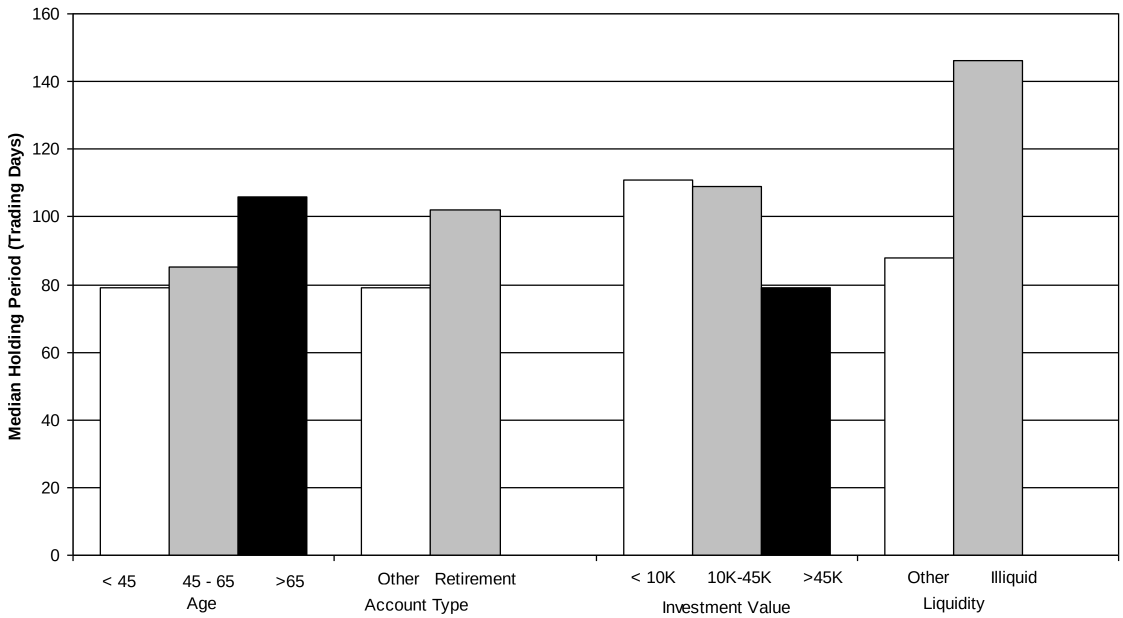


Figure 3: Survival Probabilities

This figure plots Kaplan-Meier survival probabilities for two groups of stocks held by households in the dataset. Illiquid stocks in the figure are stocks that are in the highest illiquidity decile of stocks ranked according to the adjusted Amihud illiquidity measure.

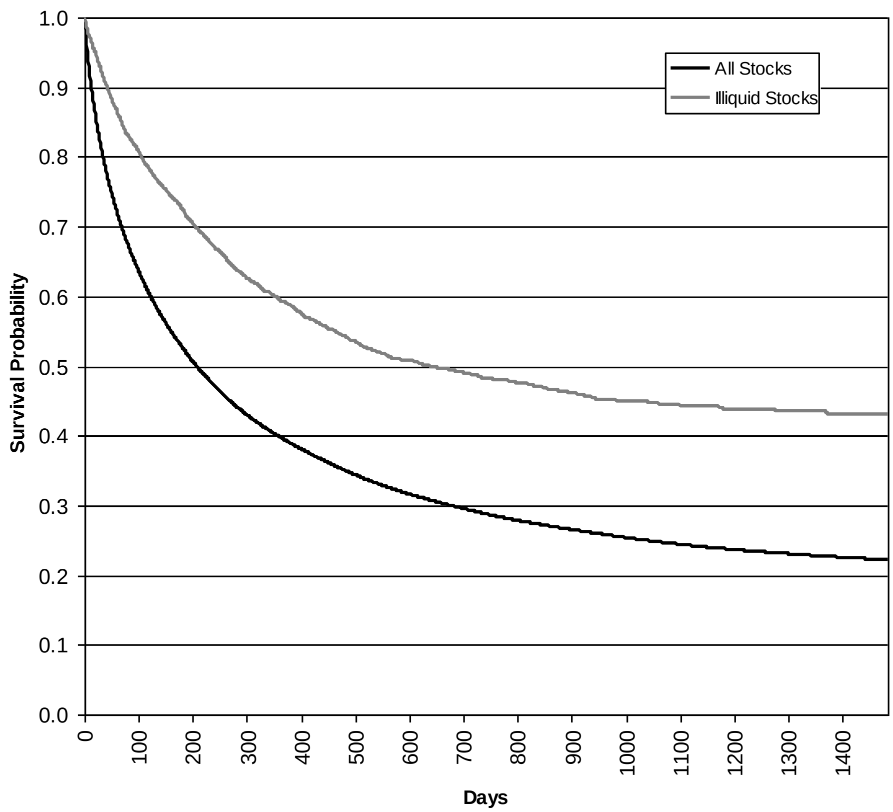


Figure 4: Hazard Ratios by Investor Sophistication

This figure plots the hazard ratios on the AdjIlliq Dum variable for different groups of investors ranked by sophistication. Hazard ratios are calculated by running a separate regression for each group of investors who have the same Sophistication value. The regression model used is the same as in Model I in Table 3.

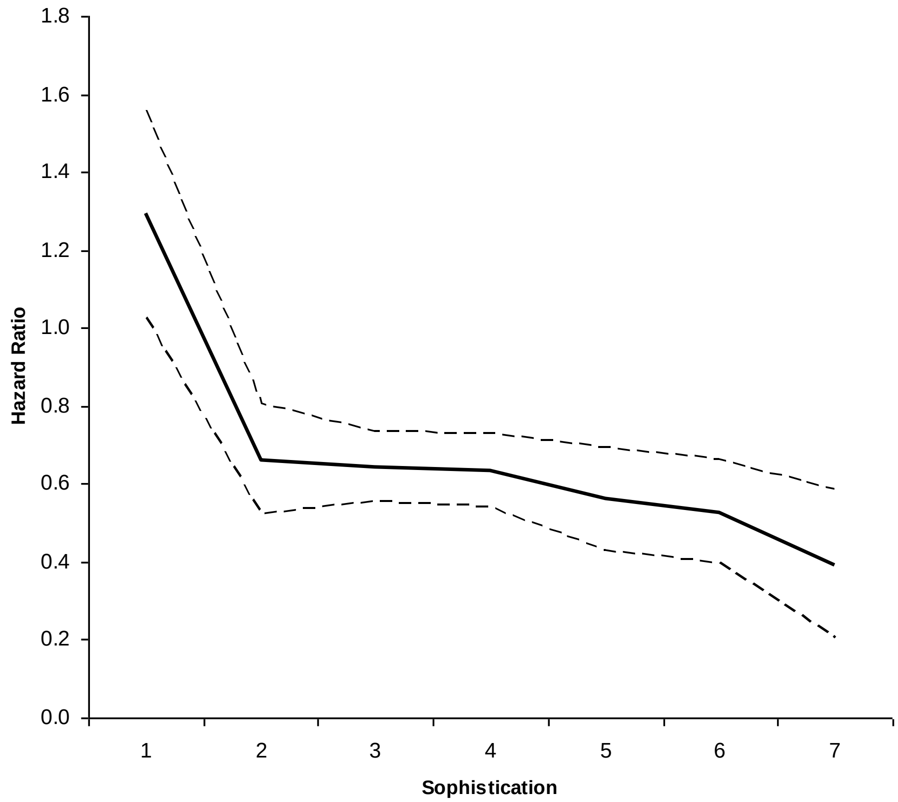


Figure 5: Distribution of Holding Periods

This figure plots the distribution of holding periods for the households in the dataset. Holding period is calculated as the average holding period for all the transactions of a given household. Positions that are not closed-out by the end of the sample period are treated as censored observations. A censored average is calculated assuming a Weibull distribution for the holding period. The figure shows distribution of holding periods calculated using positions that are closed out by the end of the sample period ('Closed' line), and calculated using censored observations ('Censored' line).

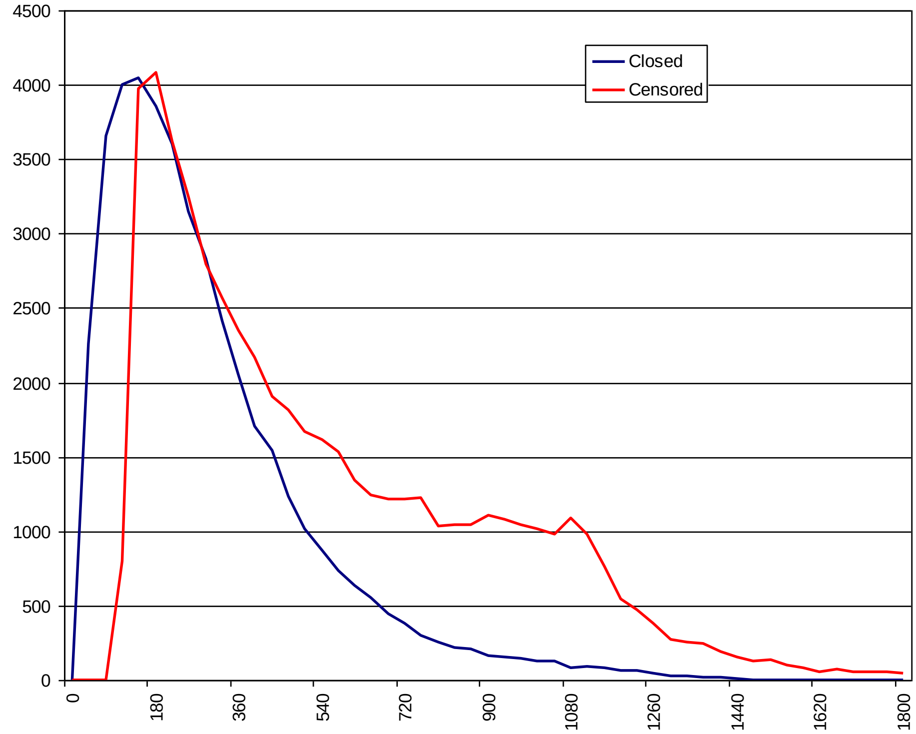


Figure 6: BSI and Illiquidity BSI

This figure plots the difference in the illiquidity ranks of buys and sells (IlliqBSI), and the aggregate level of market illiquidity (Mktilliq). Market illiquidity is calculated as the equal-weighted average of the adjusted Amihud illiquidity ratio of all stocks in a given month.