Essays in Asset Pricing

by

Deniz Anginer

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2010

Doctoral Committee:

Professor Hasan Nejat Seyhun, Chair
Professor Jeffrey Andrew Smith
Associate Professor Tyler G. Shumway
Associate Professor Uday Rajan
Assistant Professor Mario Macis
Assistant Professor Paolo Pasquariello

© Deniz Anginer

2010

For my parents, Halit and Gonul Anginer

Acknowledgements

This dissertation would not have been possible without the help of my teachers, family and friends. I thank the members of my dissertation committee, Tyler Shumway, Paolo Pasquariello, Uday Rajan, Mario Macis, Jeff Smith for generously donating their time and providing helpful comments and suggestions. I am especially grateful to my advisor, Nejat Seyhun, for his help and guidance.

I benefited greatly from numerous interactions with colleagues at the University of Michigan. I have collaborated on a number of projects with Celim Yildizhan, a fellow student, who has since become a good friend. The second chapter of my dissertation is written jointly with Celim.

Finally I thank my family for their support over the past five years. I am grateful to Melissa Panjer, without whose love and support, it would not have been possible to finish this journey.

Table of Contents

List of Tables

Table

List of Figures

Figure

Chapter I

Introduction

This work consists of three essays that investigate the effect of investor behavior on asset prices. In the first essay, titled "*Transaction Costs and Investment Decisions of Individual Investors,*" I study the liquidity decisions of 66,000 households from a large discount brokerage. My paper provides an empirical link between investors' optimal trading decisions and the liquidity premium observed in the market. In particular, I show that transaction costs are an important determinant of investors' holding periods which determine how transaction costs are amortized and priced in asset returns. I also 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.

The second essay, *"Is there a Distress Risk Anomaly? Bond Spreads as a Proxy for Default Risk,"* investigates the pricing of default risk in stock returns. The results show that credit spreads predict corporate defaults better than previously used measures, such as, bond ratings, accounting variables and structural model parameters. Contrary to previous findings, using corporate credit spreads to proxy for default risk, this study finds no significant pricing of default risk in the cross-section of equity returns. Exposure to market volatility innovations is shown to explain much of the returns to distressed stocks previously documented.

The final essay, *"Affect in a Behavioral Asset Pricing Model"*, investigates the role of psychological heuristic *Affect* in asset pricing. The paper outlines a behavioral asset pricing model where expected returns are high when objective risk is high and also when subjective risk is high. High subjective risk comes with negative affect. Investors prefer stocks with positive affect and their preference boosts the prices of such stocks and depresses their returns. Empirical support for the model is provided by studying the preferences of investors as reflected in surveys conducted by Fortune magazine during 1983- 2006. The returns of admired stocks, those highly rated by the Fortune respondents, were lower than the returns of despised stocks, those rated low. This is consistent with the hypothesis that stocks with negative affect have high subjective risk and their extra returns compensate for that risk.

Chapter II

Transaction Costs and Investment Decisions of Individual Investors

Theoretical papers link the liquidity premium to the optimal trading decisions of investors facing transaction costs. 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 drastically 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 from a large discount brokerage who make over two million trades 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 systemic 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 various household and stock characteristics. However, the effect of transaction costs on holding periods is much less than the effect predicted in the models of Vayanos (1998) and Constantinides (1988). The results in this paper offer an explanation for the discrepancy between empirically observed liquidity premium and the one predicted by these models in which the holding period is endogenously determined.[1] I find that there are differences across households in how much attention they pay to the liquidity of the securities they trade and hold in their portfolios. Investors who are more sophisticated tend to pay more attention to liquidity and have holding periods that are strongly correlated with measures of transaction costs.

There are important implications of households' liquidity decisions 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

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

postulate that investors with longer holding periods earn rents for holding illiquid securities that exceed amortized transaction costs, which drive 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, that is, households that do not pay attention to liquidity, earn lower *gross* and *net* returns.

Households tend to demand liquid securities in tandem. That is, there is systemic 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 buying liquid securities and selling illiquid securities. However, there is a subset of investors with deep pockets, those with higher incomes and higher levels of wealth, who buy illiquid securities when there is a negative liquidity shock and earn a premium in the process.

How investors make decisions in the presence of transaction costs is important not only to better understand how liquidity is priced in the financial markets, but it also has implications for investor welfare and public policy. This paper shows that expected holding periods and amortized transaction costs strongly impact the performance of household portfolios. 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 the increasingly popular notion that individual investors overtrade and lose 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, and investors who trade more frequently pay greater attention to the liquidity of the underlying stocks they trade. A number of papers also document that a subset of retail investors displays financial sophistication and market understanding and earns 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 market where a handful of investors trade to provide liquidity. Second, and more

[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.

importantly, holding periods are based on trading decisions of investors, who consider ex-ante transaction costs of the underlying securities. Another closely-related paper (Naes and Odegaard 2008) uses transaction-level Norwegian data to show that turnover is indeed a poor proxy for actual holding periods of investors.[4] 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 2.2 describes the liquidity measures and the individual trade data used herein. Sections 2.3 to 2.5 present and discuss the main findings, and Section 2.6 concludes.

2.1 Hypotheses and Related Literature

Although empirical studies document that 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.[5] 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.

[4] This research was conducted concurrently.

[5] 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).

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 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 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 priced in asset returns.

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 it is the magnitude of the relationship between holding periods and transaction costs that determines the liquidity premium in the market.[6] 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. The first hypothesis is thus:

H1a: *Holding periods are positively related to measures of fixed transaction costs after controlling for investor and stock characteristics.*

In testing the relationship between holding periods and transaction costs, I control for a number of investor and stock characteristics. I also control for the well known behavioral tendency to hang on to losing stocks too long and to sell winning stocks too quickly (the disposition effect), and the level of information asymmetry for a given stock.[7] I repeat the same analysis using portfolios instead of transactions. That is, I examine the relationship between households' overall portfolio liquidity and their average holding period. 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).

Previous studies have shown that, on average, households' stock investments perform poorly. Odean (1999), for instance, reports that individual investors' purchases under-

[6] Although in this paper 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).
[7] Asymmetric information is also considered part of transaction costs. See discussion in Section 4.

perform their sales by a significant margin.[8] However, other studies have concluded that there exists a subset of retail investors who display financial sophistication and market understanding. 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 abnormal profits. Using the same dataset in 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 the 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.*

I assume, as in Goetzmann and Kumar (2008), that the level of financial sophistication is correlated with education and resources available to an investor. I use income, wealth invested in the stock market and the occupation of the investor to proxy for financial sophistication. I also use information contained in investors' trades. I assume that

[8] Barber and Odean (2000, 2001), using the same dataset in 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.

investors who engage in short selling, who trade options or who trade foreign securities are likely to be more sophisticated than the average investor.

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.[9] 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 (rents) when investing in that security because their amortized transaction costs are lower, which imply:

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

[9] Vayanos and Vila (1997) show a similar result when securities are identical except for transaction costs.

could, therefore, also indicate lack of financial sophistication and market knowledge, which is associated with lower gross returns:

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 them having 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, as to what causes this common variation. Commonality in liquidity can arise from the supply side, if there is systemic variation in the costs of providing liquidity.[10] Commonality can also arise from the demand side, if a common factor such as volatility or uncertainty causes a systemic variation in the demand for liquidity.[11] 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 market volatility affect systemic liquidity by creating correlated trading patterns among

---

[10] 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.

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

investors. By examining the actual trades of investors, I can test whether there is systemic variation in the demand for liquid assets and whether liquidity shocks apply (or transmitted) systemically across investors that can potentially cause market-wide effects:

H3a: *There is systemic variation in households' demand for liquid stocks.*

In order to test whether there is systemic variation in the demand for liquidity, I employ a similar methodology used in Kumar and Lee (2006) and Barber, Odean and Zhu (2003), who investigate systemic correlation in the trades of individual investors. I test whether randomly selected non-overlapping groups of investors tilt their portfolios towards liquid assets at the same time.

If there is systemic variation in demand for liquid assets across investors, it is important to examine how this systemic 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 providing 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 retail traders providing 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.[12] 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, and test if households are net demanders or suppliers of liquid securities when aggregate market liquidity is low:

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

The recent Goldman Sachs' agreement to sell $5 billion of perpetual preferred stock to Berkshire Hathaway illustrates both the adverse effects of market participants seeking liquidity at the same time and the importance of external investors with deep pockets as liquidity providers. 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. 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 suppliers of illiquid stocks when aggregate market liquidity is low.*

To test the above hypothesis, I construct an aggregate market liquidity measure as in Acharya and Pedersen (2005) and Amihud (2002), and compare the liquidity levels of purchases and sales of stocks by households under different liquidity regimes. I use a

[12] In most of these studies, investors can not be identified and their transactions can not be tracked over time.

regression framework to test the effect of investor characteristics, such as income and wealth, on the demand for liquid securities.

2.2 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-dimensional 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-dimensional 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:

$$c_{i,t} = \begin{cases} \sqrt{-\operatorname{cov}(r_{i,t}, r_{i,t-1})} & \text{if } \operatorname{cov}(r_{i,t}, r_{i,t-1}) < 0; \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

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 (2.1) 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).[13]

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

$$Illiq_{i,t} = \frac{1}{D_{i,t}} \sum_{d=1}^{D_{i,t}} \frac{\left|r_{i,d}\right|}{dvol_{i,d}} \tag{2.2}$$

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

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

impact of a trade. I adjust this measure as in Acharya and Pedersen (2005) to make it stationary and to remove outliers:

$$AdjIlliq_{i,t} = \min\left(0.25 + 0.30 \times Illiq_{i,t} \times M_{t-1}, 30\right) \tag{2.3}$$

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:

$$r^e_{i,d+1,t} = \theta_{i,t} + \phi_{i,t} r_{i,d+1,t} + \gamma_{i,t} \text{sign}(r^e_{i,d,t}) v_{i,d,t} + \varepsilon_{i,d,t} \tag{2.4}$$

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):

$$\begin{aligned} \mathrm{SprBuy} &= \frac{P_{crsp}}{P_{buy}} - 1 \\ \mathrm{SprSell} &= 1 - \frac{P_{crsp}}{P_{sell}} \end{aligned} \tag{2.5}$$

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.[14] 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 2.5. Table 2.1 reports the summary statistics and correlations for the liquidity measures for stocks traded by households in the dataset.

2.3 Holding Periods and Transaction Costs

2.3.1 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.[15] 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 underlying stocks' liquidity.[16] The median holding period is shorter for stocks held in retirement accounts. Investors who are older and who have less wealth invested in the

[14] In the analyses that follow, I also explicitly control for other potential determinants of holding periods such as stock and investor characteristics.
[15] 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).
[16] 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.

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 2.2, 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.[17] With hazard models, an investor's trade decision can be explicitly modeled by considering the investor's sell-hold decision each day. In this paper, I use a Cox proportional hazard model with potentially time varying explanatory variables.[18] The hazard model takes the form:

$$\lambda(t) = \lambda_0(t)\exp\left(x(t)'\beta + z'\alpha\right) \tag{2.6}$$

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 constant or time varying. In equation (2.6), *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

[17] 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.
[18] Details about estimating the proportional hazard model can be found in Cox and Oakes (1984).

purchase price. This dummy variable acts as a proxy for the disposition effect. 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 that follow.[19] In the tables that follow, 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.

Table 2.3 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 the12 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

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

not belonging to that group.[20] 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 2.3 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 2.3), those who are sophisticated tend to pay attention to the transaction 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

[20] 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.

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. *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 2.4 describes the criteria used to construct the *Sophistication* variable. I run the same hazard regression as before (Model I in Table 2.3), 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 2.4, 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).

2.3.2 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 2.5 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 2.5. 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 (Garleanu and Pedersen 2007) could reduce the correlation between holding periods and liquidity. 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.[21] 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 2.5 under Model V. The PIN measure significantly reduces the holding period of investors. 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

[21] 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.

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 2.5 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.

2.3.3 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. Specifically, 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:

$$PIlliq_{i,t} = \frac{\sum_{k=1}^{N} \frac{AdjIlliq_{i,t}^{k}}{MktIlliq_{t}} \times \left| Eq_{i,t}^{k} \right|}{\sum_{k=1}^{N} \left| Eq_{i,t}^{k} \right|} \tag{2.7}$$

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.[22] 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 2.6 shows the results from regressing average portfolio liquidity on household holding periods and household characteristics:

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

$$PIlliq_i = \beta_0 + \beta_{HP} HP_i + \sum_{k=1}^{K} \beta_k InvCh_{i,k} \tag{2.8}$$

In equation (2.8), $PIlliq_i$ 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 2.6. Holding period is a statistically significant determinant of portfolio liquidity. Given that the median and the 75^{th} 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 with hypothesis $H1a$.

2.4 Holding Periods and Returns

2.4.1 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 2.1. 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.[23] To be able to make comparisons across different holding periods, I calculate average daily returns from cumulative raw and excess returns as:

$$avgr_i = \sqrt[1/HP]{\prod_{d=1}^{HP}\left(1 + r_{i,d}\right)} - 1 \qquad (2.9)$$

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 2.2. 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

[23] 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.

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 2.7.[24] 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 2.1, in the 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 and not returns for the whole portfolio, the results could be biased if only profiTable 2.trades are closed out producing a disposition

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

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. In response, I consider returns for fixed holding periods from the day of purchase (1, 6, 12 month returns are also reported in Table 2.7). 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 2.3. I eliminate households with a positive *disp* coefficient calculated with a 10% confidence level or higher. 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 2.6. 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*.

2.4.2 Liquidity Decisions and Returns

There are cross-sectional differences in the correlation between holding periods and transaction costs across households. As described in Section 2.1, 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. 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.[25] The summary statistics for the transaction costs coefficient calculated from household level hazard regressions are reported in Table 2.8. 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 2.9. 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

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

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.

2.5 Individual Investors and Demand for Liquid Securities

2.5.1 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 2.1, 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 systemic variation in the demand for 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:

$$IlliqBSI_t^G = \frac{\sum_{i \in G} V_{t,Buy}^i \times AdjIlliqRank_t^i - \sum_{i \in G} V_{t,Sell}^i \times AdjIlliqRank_t^i}{\sum_{i \in G} V_{t,Buy}^i \times AdjIlliqRank_t^i + \sum_{i \in G} V_{t,Sell}^i \times AdjIlliqRank_t^i} \tag{2.10}$$

where $V_{t,Buy}^i$ and $V_{t,Sell}^i$ are the total value of buys and sells, respectively, for investor *i* in month *t*. $AdjIlliRank_t$ is the weighted average adjusted illiquidity rank of stock holdings of investor *i* belonging to group *G* in month *t* 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. If the demand for liquid securities is independent 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:

$$IlliqBSI_t^G = \beta_0^G + \beta_{MKT}^G MKT_t + \beta_{BSI}^G BSI_t + \varepsilon_t^G \tag{2.11}$$

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:

$$BSI_t = \frac{\sum_{i \in N} V^i_{t,Buy} - \sum_{i \in N} V^i_{t,Sell}}{\sum_{i \in N} V^i_{t,Buy} - \sum_{i \in N} V^i_{t,Sell}} \tag{2.12}$$

$V^i_{t,Buy}$ and $V^i_{t,Sell}$ 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^G_t, for different pairs of investor groups.

The results are reported in Table 2.10. 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 systemic component in the demand for liquid securities across households. The results support hypothesis *H3a*, that there is systemic variation in the demand for liquid securities.

2.5.2 Aggregate Market Liquidity and Household Demand for Liquid Securities

As mentioned in Section 2.1, a number of papers treat retail investors as noise traders providing constant liquidity to the market. However, if there is systemic 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 systemic 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).[26] 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:

$$IlliqBSI_t^{ALL} = \frac{\sum_{i \in N} V_{t,Buy}^i \times AdjIlliqRank_t^i - \sum_{i \in N} V_{t,Sell}^i \times AdjIlliqRank_t^i}{\sum_{i \in N} V_{t,Buy}^i \times AdjIlliqRank_t^i + \sum_{i \in N} V_{t,Sell}^i \times AdjIlliqRank_t^i} \qquad (2.13)$$

$V_{t,Buy}^i$ and $V_{t,Sell}^i$ are the total value of buys and sells, respectively, for investor *i* in month *t*. $AdjIlliqRank_t^i$ is the weighted average adjusted illiquidity rank of stock holdings of investor *i* in month *t* using one month lagged adjusted illiquidity ranks. I compute the sum over all *N* investors. 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. Consistent with the previous studies, I find that there are more buys when market illiquidity is high. However, once we consider the liquidity level of the underlying stocks that are traded, the picture

[26] 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%.

changes. 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 2.10 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 8^{th} 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 2.11 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:

$$TransAdjIlliqrank_{k,t} = \beta_0 + \beta_1 Buy_{k,t} + \beta_2 Affluent_i + \beta_3 MktIlliqDum_t$$
$$+\beta_4 Buy_{k,t} \times Affluent_i + \beta_5 Buy_{k,t} \times MktIlliqDum_t + \beta_6 MktIlliqDum_t \times Affluent_i \quad (2.14)$$
$$+\beta_7 Buy_{k,t} \times MktIlliqDum_t \times Affluent_i$$

In equation (2.14), $TransAdjIlliqrank_{k,t}$ is the lagged adjusted illiquidity rank of the underlying stock for transaction *k* in month *t*.[27] 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 invested more than $100,000 in the stock market during the sample period.[28] Model I in Table 2.11 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

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

interaction term, $MktIlliq\ \ Dum \times Buy \times Affluent$, in Model IV in Table 2.11 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.

2.6 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. 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 systemic 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.

Table 2.1: 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 2.2: 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 2.3: 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 2.4: 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*
Month Dummies	Yes	

Table 2.5: 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 VII**	**Model IX**
AdjIlliq Dum	0.587***						0.660***	0.687***	1
	<.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
						<.0001			*<.0001*
AdjIlliq Dum * Disp Dum									1
									<.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 2.6: 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 2.3. 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 2.7: 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 2.8: 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 2.5. *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 2.9: 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 2.5. *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 2.10: Common Demand for Liquidity

This table reports correlation statistics from three different simulations that test for a systemic 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 2.11: 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*

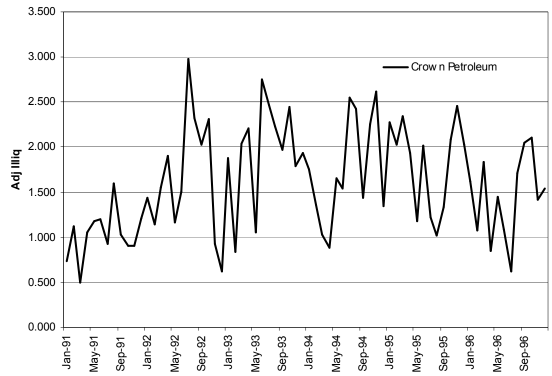
Table 2.12: Market Liquidity and Liquidity of Transactions

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 2.1: Illiquidity Ratio

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



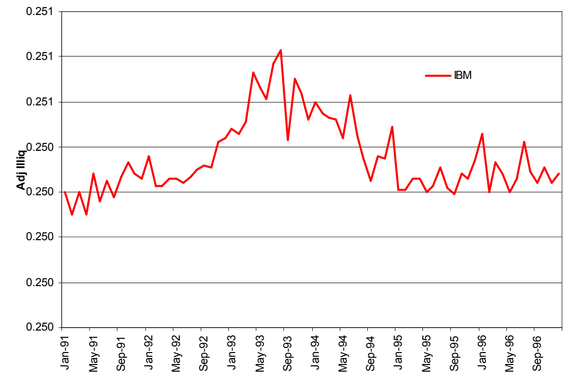


Figure 2.2: Holding Periods of Households

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.

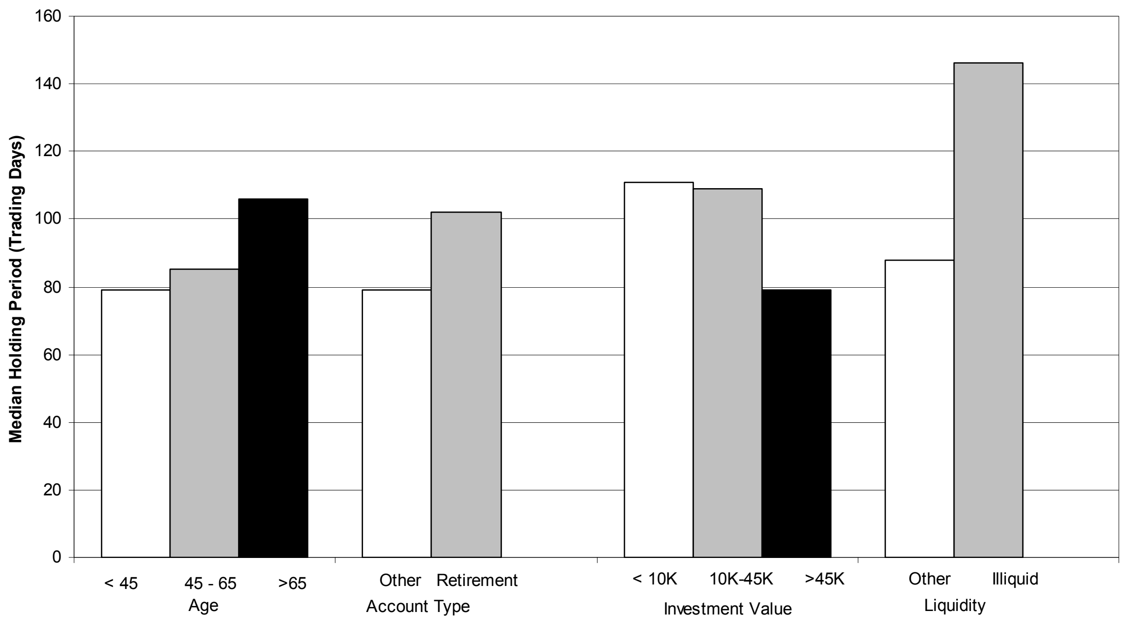


Figure 2.3: Survival Probabilities

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.

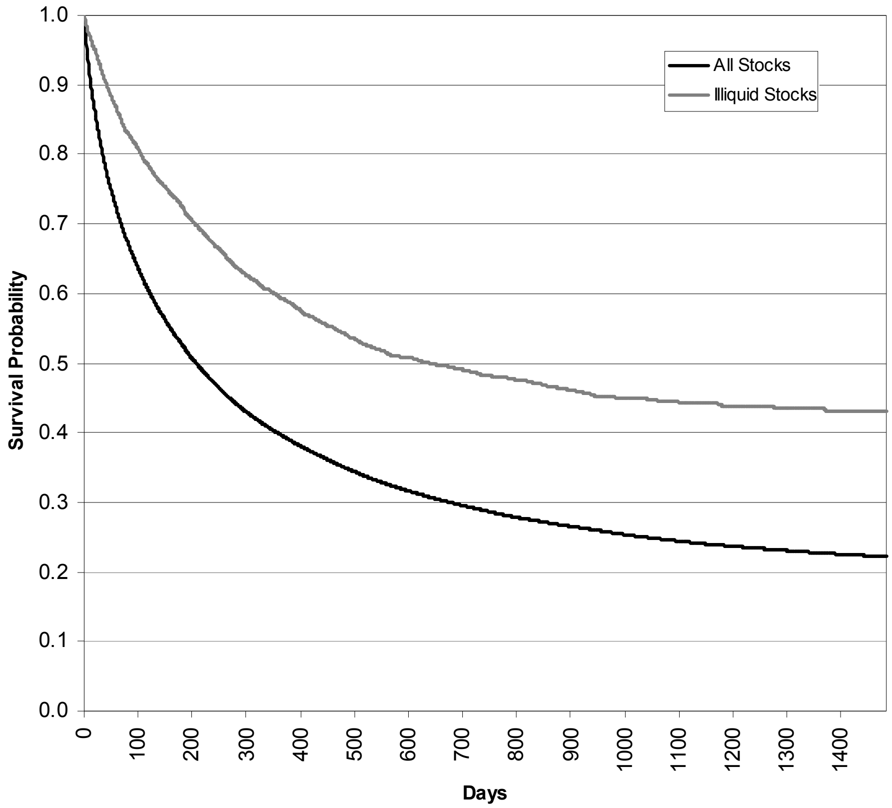


Figure 2.4: Hazard Ratios by Investor Sophistication

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.

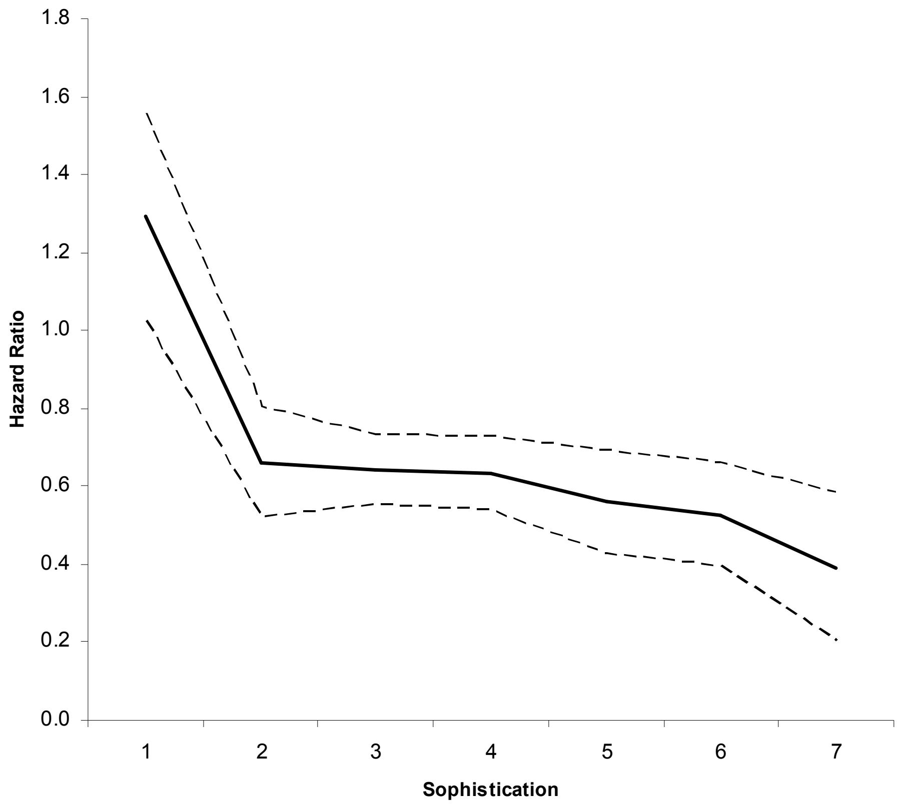


Figure 2.5: Distribution of Holding Periods

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).

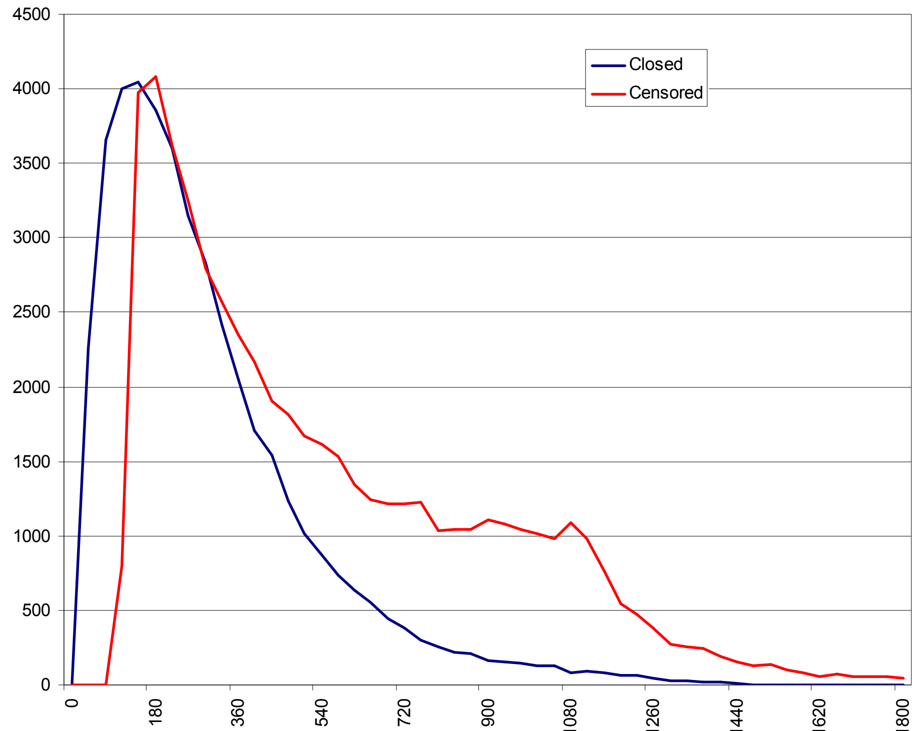
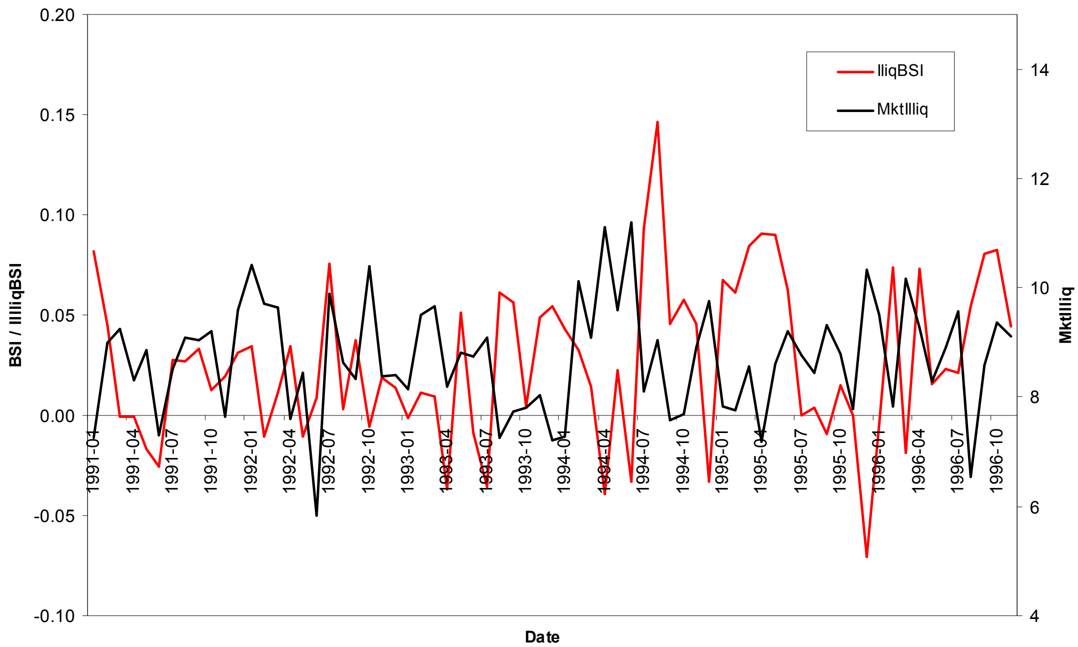


Figure 2.6: BSI and Illiquidity BSI

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.



Chapter III

Is there a Distress Risk Anomaly?

Corporate Bond Spread as a Proxy for Default Risk

A fundamental tenet of asset pricing is that investors should be compensated with higher returns for bearing systematic risk that can not be diversified. Recently a number of papers examined whether default risk is such a systematic risk and whether it is priced in the cross section of equity returns. On the theoretical side, default risk can be a priced factor if a firm's *Beta* within the framework of the Capital Asset Pricing Model (CAPM) does not fully capture default-related risk. Default risk may not be fully correlated with the market itself, but could be related to declines in other un-measured components of wealth such as human capital (Fama and French 1996) or risk related to debt securities (Ferguson and Shockley 2003), distinct from risk related to equities. Empirically, research thus far has focused on determining the ex-ante probability of firms failing to meet their financial obligations and testing to see if there is co-movement in security returns of firms in response to changes in an empirically constructed default risk factor. Previous studies have utilized different proxies and approaches to measure financial distress and have found anomalously low returns for stocks with high probabilities of default.[29] The low returns on stocks with high default risk cannot be explained by Fama and French (1993) risk factors. Stocks with high distress risk tend to have higher market betas and load more heavily on size and value factors leading to significantly negative alphas.

In this paper we argue that the anomalous results documented in the literature are due to the poor quality of the proxies used to measure default risk. First, previous papers measure financial distress by determining the ex-ante real-world probability of default, as opposed to risk-neutral probability of default that incorporates a risk premium for systematic risk.[30] Ranking stocks on their real-world default probabilities, as done in Dichev (1998) and Campbell, Hilscher and Szilagyi (2008), implicitly assumes that high default probability stocks also have high exposure to the systematic component of default risk. George and Hwang (2009) show that a firm's ex-ante real-world probability of default does not necessarily reflect the firm's exposure to systematic default risk. Furthermore, it has been well documented (see for instance Almeida and Philippon 2007 and Berndt et al. 2005) that there is a substantial difference between the risk-neutral and historical probabilities of default. Second, previous papers have shown three stock characteristics – idiosyncratic volatility, leverage and profitability – to be most closely associated with high corporate default rates. High idiosyncratic volatility, high leverage and low profitability predict high default probability. However these are the same characteristics that are known to be associated with future expected returns. Within the q-theory framework (Cochrane 1991, Liu, Whited and Zhang 2008), low profitability (more

[29] See for instance Dichev (1998) and Campbell, Hilscher and Szilagyi (2008).
[30] Almeida and Philippon (2007), Hull, Predescu and White (2006) provide empirical evidence on the difference between real-world and risk-neutral default probabilities implied by credit spreads.

likely to default) firms have low expected future returns. Similarly, firms with high leverage (more likely to default) and high idiosyncratic volatility (more likely to default) have low stock returns (Penman et al. 2007, Ang, Hodrick, Xing and Zhang 2008). It is not clear if the distress anomaly is just the manifestation of one or more of these previously documented return relationships. We show that the difference in returns between high and low distress stock portfolios becomes insignificant once we control for these three stock characteristics.

In this paper, we take a different approach to measuring default risk and use a market based measure, namely corporate bond spreads, to proxy for distress risk. This measure offers several advantages over others that have been utilized in the literature thus far. Unlike structural models of corporate bankruptcy that make simplifying assumptions about the capital structure of a firm, our proposed measure is model and assumption free. And unlike stock characteristics used to measure default risk, which may reflect information about future returns unrelated to distress risk, credit spreads reflect the market consensus view of the credit quality of the underlying firm. Moreover, credit spreads contain a risk-premium for systematic risk. As such, unlike previously used measures, credit spread, is a proxy for the market-implied risk-adjusted (or risk-neutral) probability of default and is a more complete measure of default risk. We show that credit spreads predict corporate defaults better than previously used measures based on structural models, bond ratings and accounting variables. Using this market based measure, we find that there is no evidence of firms with high default risk delivering anomalously low returns, and we do not find default risk to be a priced risk factor in the cross-section of equity returns.

Ours is not the first paper to study the relationship between default risk and equity returns. Dichev (1998) uses Altman's z-score and Ohlson's o-score to measure financial distress. He finds a negative relationship between default risk and equity returns during the 1981-1995 time period. In a related study, Griffin and Lemmon (2002), using the o-score to measure default risk, find that growth stocks with high probabilities of default have low returns. Using a comprehensive set of accounting measures, Campbell, Hilscher and Szilagyi (2008) (hereafter CHS) show that stocks with high risk of default deliver anomalously low returns. Garlappi, Shu, and Yan (2008), who obtain default risk measures from Moody's KMV, also find similar results to those of Dichev (1998) and CHS (2008). They attribute their findings to the violation of the absolute priority rule.

George and Hwang (2009) argue that the negative relation between returns and leverage can explain the pricing of distress risk anomaly. Avramov et al. (2009) show that most of the negative return for high default risk stocks is concentrated around rating downgrades. Vassalou and Xing (2004) find some evidence that distressed stocks, mainly in the small value group, earn higher returns.[31] Chava and Purnanandam (2008) argue that the poor performance of high distress stocks is limited to the post-1980 period when investors were positively surprised by defaults. When they use implied cost of capital estimates from analysts' forecasts to proxy for ex-ante expected returns, they find a positive relation between default risk and expected returns.

[31] Da and Gao (2005) argue that Vassalou and Xing's results are limited to one month returns on stocks in the highest default likelihood group which trade at very low prices. They show that returns are contaminated by microstructure noise and the positive one month return is compensation for increased liquidity risk.

Our paper also contributes to a growing literature on bankruptcy prediction.[32] In particular, we show the importance of market based variables in predicting bankruptcy. Corporate bond spreads significantly increase the pseudo R^2's in hazard regressions when we run a horse race of corporate spreads with a comprehensive set of accounting measures, bond ratings and structural model parameters previously used in the literature. Adding corporate spread to the covariates used in CHS (2008), for instance, increases the pseudo R^2 from 27.6% to 37.4%.[33] These results strongly indicate that corporate bond spreads contain default information above and beyond the measures commonly used in the literature.

The rest of the paper is organized as follows. Sections 3.1 and 3.2 describe the data and the different default measures used in this study. Section 3.3 reports the return analyses for high default risk stocks and examines the relationship between various stock characteristics and default risk. Section 3.4 describes the use of credit spreads as a predictor of corporate bankruptcy and as a proxy for default risk, and also contains the asset pricing tests to see if default risk, as measured by credit spreads, is priced in the cross section of equity returns. Section 3.5 concludes.

3.1 Data

In this section, we briefly describe the data sources used in this study. Firm level accounting and price information are obtained from COMPUSTAT and CRSP for the 1980–2008 time period. We exclude financial firms (SIC codes 6000 through 6999) from

---

[32] See for instance Altman (1968), Zmijewski (1984), Ohlson (1986), Shumway (2001), and Chava and Jarrow (2004).

[33] Using corporate spread as the lone predictor variable yields a pseudo R^2 of 26.5%, similar to the pseudo R^2 obtained from using all of the CHS (2008) covariates.

the sample. To avoid the influence of microstructure noise we also exclude firms priced less than one dollar in the analyses that follow. The data items used to construct distress measures are explained in detail in the Appendix.

Corporate defaults between 1981 and 2008 are identified from the Moody's Default Risk Services' Corporate Default database, SDC Platinum's Corporate Restructurings database, Lynn M. LoPucki's Bankruptcy Research Database, and Shumway's (2001) list of bankruptcies. We choose 1981 as the earliest year for identifying bankruptcy filings as the Bankruptcy Reform Act of 1978 is likely to have caused the associations between accounting variables and the probability of bankruptcy to change. Furthermore, we have little corporate bond yield information prior to 1980. In all, we obtain a total of 548 firm defaults covering the period 1981–2008, for which we have complete accounting-based measures. 94 of these bankruptcies also have corresponding corporate bond spread information.

Corporate bond data used in this study comes from three separate databases: the Lehman Brothers Fixed Income Database (Lehman) for the period 1974 to 1997, the Fixed Investment Securities Database (FISD) for the period 1998 to 2002, and the Trade Reporting and Compliance Engine (TRACE) system dataset from 2003 to 2008. We also use the National Association of Insurance Commissioners Database (NAIC) for bond descriptions. Due to the small number of observations prior to the year 1980, we include only the period 1980 to 2008 in the analyses that follow.

Our sample includes all U.S. corporate bonds listed in the above datasets that satisfy a set of selection criteria commonly used in the corporate bond literature.[34] We exclude all

[34] See for instance Duffee (1999), Collin-Dufresne, Goldstein, and Martin (2001) and Avramov et al. (2006).

bonds that are matrix-priced (rather than market-priced) from the sample. We remove all bonds with equity or derivative features (i.e. callable, puttable, and convertible bonds), bonds with warrants, and bonds with floating interest rates. Finally, we eliminate all bonds that have less than one year to maturity.

For all selected bonds, we extract beginning of month credit spreads calculated as the difference between the corporate bond yield and the corresponding maturity matched treasury rate. There are a number of extreme observations for the variables constructed from the different bond datasets. To ensure that statistical results are not heavily influenced by outliers, we set all observations higher than the 99th percentile value of a given variable to the 99th percentile value. All values lower than the first percentile of each variable are winsorized in the same manner. For each firm, we calculate a value-weighted average of that firm's outstanding bond spreads, using market values of the bonds as weights. There are 107,692 firm months and 1011 unique firms with credit spread and firm level data. There is no potential survivorship bias in our sample as we do not exclude bonds that have gone bankrupt or those that have matured.

As not all companies issue bonds, it is important to discuss the limitations of our dataset. We compute summary statistics for default measures and financial characteristics of the companies in our bond sample and for all companies in CRSP. These results are summarized in Table 3.1. Not surprisingly, companies in the bond sample are larger and show a slight growth tilt. There is, however, significant dispersion in size, market-to-book, and credit spread values. The bond sample covers a small portion of the total number of companies, but a substantial portion in terms of total market capitalization. For instance, in the year 1997, the number of firms with active bonds in our sample constitutes about 4% of all the firms in the market. However, in terms of market capitalization, the dataset captures about 40% of aggregate equity market value in 1997. In section 3.3, we show that the distress anomaly as described by CHS (2008) and others exists in our bond sample.

3.2 Default Risk Measures

There is a vast literature on the statistical modeling of the probability of bankruptcy. In this paper, we create measures of financial distress based on three models of bankruptcy prediction that have been utilized by previous researchers investigating the pricing of distress risk.

3.2.1 Static Models

Static models of bankruptcy prediction use firm specific accounting information, employing either a multiple discriminant analysis as in Altman (1968) or a conditional logit model as in Ohlson (1980), in order to assess which firm characteristics are important in determining the probability of financial distress. These models then use the estimates from the single period classification to predict future implied probability of bankruptcy.[35] In this paper, we use Altman's z-score and Ohlson's o-score, two popular frameworks that have been widely used in empirical research and practice. Altman's z-score is defined as the following:

$$\text{z-score} = 1.2\ WCTA + 1.4\ RETA + 3.3\ EBITTA + 0.6\ METL + 1.0\ STA \qquad (3.1)$$

[35] Using single period observations introduce a bias in static models as discussed in Shumway (2001).

where $WCTA$ is the ratio of working capital to total assets, $RETA$ is the ratio of retained earnings to total assets, $EBITTA$ is the ratio of earnings before interest and taxes to total assets, *METL* is the ratio of market equity to total liabilities, and *STA* is the ratio of sales to total assets. Ohlson's o-score is defined as:

$$\begin{aligned} o\text{-score} &= -1.32 - 0.407\log(SIZE) + 6.03\ TLTA - 1.43\ WCTA \\ &+0.076\ CLCA - 1.72\ OENEG - 2.37\ NITA - 1.83\ FUTL \\ &+0.285\ INTWO - 0.521\ CHIN \end{aligned} \tag{3.2}$$

where *SIZE* is total assets divided by the consumer price index, *TLTA* is the ratio of total liabilities to total assets, $CLCA$ is the ratio of current liabilities to current assets, *OENEG* is a dummy variable equal to one if total liabilities exceeds total assets and zero otherwise, *NITA* is the ratio of net income to total assets, *FUTL* is the ratio of funds from operations to total liabilities, *INTWO* is a dummy variable equal to one if net income was negative for the past two years and zero otherwise, and *CHIN* is a measure of the change in net income. The accounting variables used to construct the z-score and the o-score are described in detail in the appendix.

3.2.2 Dynamic Models

Dynamic models of bankruptcy prediction (Shumway 2001, Chava and Jarrow 2004 and CHS 2008) use a dynamic panel model approach and incorporate market based variables such as market capitalization and past equity returns. Dynamic models of bankruptcy prediction avoid the biases of the static models by adjusting for potential duration dependence issues. In this paper we use the CHS (2008) specification:

$$
\begin{aligned}
CHS\text{-}score_t \;=\;& -9.164 - 20.264\ NIMTAAVG_t + 1.416\ TLMTA_t \\
& -7.129\ EXRETAVG_t + 1.411\ SIGMA_t - 0.045\ RSIZE_t \\
& -2.132\ CASHMTA_t + 0.075\ MB_t - 0.058\ PRICE_t
\end{aligned} \tag{3.3}
$$

where *NIMTAAVG* is a geometrically declining average of past values of the ratio of net income to the market value of total assets, *TLMTA* is the ratio of total liabilities to the market value of total assets, *EXRETAVG* is a geometrically declining average of monthly log excess stock returns relative to the S&P 500 index, *SIGMA* is the standard deviation of daily stock returns over the previous three months, *RSIZE* is the log ratio of market capitalization to the market value of the S&P 500 index, *CASHMTA* is the ratio of cash to the market value of total assets, *MB* is the market-to-book ratio, *PRICE* is the log price per share truncated from above at $15.[36]

3.2.3 Structural Model

The third measure we use in this study is based on the structural default model of Merton (1974). This approach treats the equity value of a company as a call option on the company's assets. The probability of bankruptcy is based on the "distance-to-default" measure, which is the difference between the asset value of the firm and the face value of its debt, scaled by the standard deviation of the firm's asset value. There are a number of different approaches to calculating the distance-to-default measure. We follow CHS (2008) and Hillegeist et al. (2004) in constructing this measure, the details of which are provided in the appendix.

[36] In computing the CHS-score, we use coefficients on the variables calculated from rolling regressions to avoid a look-ahead bias. We thank Jens Hilscher for providing this data.

3.3 Pricing of Default Risk

3.3.1 Returns to Distressed Stocks

In this section we analyze the effect of default risk on stock returns. We sort stocks into deciles each January from 1981 through 2008, according to their default probabilities calculated using the CHS hazard model, Ohlson's o-score, and Merton's distance-to-default measure.[37] In the analyses that follow, we exclude financial firms (SIC codes 6000 through 6999); we also exclude firms priced less than one dollar as of the portfolio formation date from the sample to avoid the influence of microstructure noise. The stocks in each decile portfolio are held for a year. Following CHS (2008), if a delisting return is available we use the delisting return, otherwise we use the last available return in CRSP. We repeat the same analyses for stocks in our bond dataset. To save space we only report returns for the top and bottom, and the difference between top and bottom deciles.

We compute the value-weighted return for these decile portfolios on a monthly basis and regress the portfolio return in excess of the risk-free rate on the market (*MKT*), size (*SMB*), value (*HML*), and momentum (*MOM*) factors:

$$\mathrm{r}_t^i = \alpha^i + \beta_{MKT}^i MKT_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \beta_{MOM}^i MOM_t + \varepsilon_t^i \qquad (3.4)$$

The results are reported in Table 3.2. Panels A, B, C show returns for default risk portfolios calculated using the CHS hazard model, Ohlson's o-score, and Merton's

[37] We obtain similar results using Altman's z-score, which are not reported to save space.

distance-to-default measure respectively. The results under 'Bond Sample' on the right hand side include only the companies in our bond sample.

Our results are consistent with those obtained in the previous studies. Stocks in the highest default risk portfolio have significant negative returns. Using the CHS default probability, the difference in returns between the highest and lowest default risk portfolios is -1.24% per month. The intercepts from the market and the 4-factor model are economically and statistically significant. Monthly alpha for the zero cost portfolio formed by going long on stocks in the highest default risk decile and short on stocks in the lowest default risk decile is -0.83% per month. We find similar results using Merton's distance-to-default measure (monthly 4-factor alpha equal to -0.62%), and Ohlson's o-score (monthly 4-factor alpha equal to -1.28%) to form default risk portfolios. The results are weaker for the bond sample, but still economically and statistically significant. The 4-factor monthly alphas for the high minus low zero cost default risk portfolios are -0.32%, -0.10% and -0.24% using the CHS hazard model, Ohlson's o-score, and Merton's distance-to-default measure respectively.

The loadings on the size and value factors suggest that distressed stocks are mostly small and value stocks. The loading on the momentum factor is consistent with the intuition that distressed stocks tend to have low returns prior to portfolio formation. These results are consistent across different measures of distress, and the results hold in our bond sample.

3.3.2 Stock Characteristics and Distress Returns

Previous research has identified a number of stock characteristics that predict high default probabilities for companies. However, three characteristics – leverage, idiosyncratic volatility and profitability – have been shown to be most closely associated with corporate default rates. High leverage, high idiosyncratic volatility and low profitability predict higher rates corporate default. As mentioned earlier, these are the same characteristics that are ex-ante associated with low future returns. Ang, Hodrick, Xing and Zhang (2006, 2008) establish a robust relationship between idiosyncratic volatility and stock returns. This negative relationship has been termed the 'idiosyncratic volatility puzzle', since rational asset pricing theories predict that the relationship be positive or that there be no relationship at all.[38] There have been some behavioral and agency-based explanations for the negative relationship between idiosyncratic volatility returns.[39] Penman et al. (2007) show a negative relationship between leverage and stock returns – the leverage anomaly. Similarly, low profitability predicts low returns. Q-theory provides the theoretical link between profitability and equity returns (Cochrane 1991, Liu, Whited and Zhang 2008). It is not clear if distress anomaly is just an amalgamation of one or more of these previously documented return relationships. In this section we investigate in detail the relationship between default risk and these three stock characteristics. In particular we want to see if the distress anomaly persists once we explicitly control for idiosyncratic volatility, profitability and leverage.

---

[38] Merton (1987), Malkiel and Xu (2002) and Jones and Rhodes-Kropf (2003) link higher returns on high-volatility stocks to investors not being able to diversify.

[39] The behavioral model of Barberis and Huang (2001) predicts that higher idiosyncratic volatility stocks should earn higher expected returns. Falkenstein (1996) reports that mutual fund managers prefer to hold more volatile stocks for the upside option value they provide.

To control for these three stock characteristics, we perform a double sort. We sort stocks into five groups each January from 1981 to 2008 according to the CHS probability of default. Then within each distress group we sort stocks based on the previous year's stock characteristic (idiosyncratic volatility, profitability or leverage) into five groups, creating a total of 25 portfolios. We then calculate 4-factor alphas for the distress portfolios after controlling for the effects of the characteristics. We do this by averaging the returns of the five distress portfolios over each of the characteristic portfolios. We use *NIMTAAVG* as the profitability measure and *TLMTA* as the leverage measure. Both variables are described in Section 3.2. We follow AHXZ (2006) and calculate idiosyncratic volatility relative to the Fama-French 3-factor model. First, we regress daily stock returns from the previous calendar year on the Fama-French 3 factors:

$$r_t^i = \alpha^i + \beta_{MKT}^i MKT_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \varepsilon_t^i \tag{3.5}$$

Idiosyncratic volatility is then calculated as the standard deviation of the residuals: $\sqrt{\mathrm{var}\left(\varepsilon_t^i\right)}$.

Panel A of Table 3.3 reports 4-factor alphas for the five distress portfolios, as well as 4-factor alphas for the distress portfolios after controlling for the three stock characteristics. We also report in Panel B of Table 3.3, average idiosyncratic volatility, leverage and profitability values for firms belonging to each of the five distress portfolios. There is a strong relationship between distress risk and the three stock characteristics. Idiosyncratic volatility increases monotonically from 2.5% for the lowest distress group to 4.5% for the highest group. Leverage increases from 0.22 for the lowest distress group to 0.61 for the highest distress group. Similarly, profitability for the lowest distress group is 1.2% and decreases monotonically to -1.1%. The unconditional 4-factor alpha for the zero cost portfolio formed by going long high distress stocks and shorting low distress stocks is -0.88% per month, yet this premium decreases to -0.61% after controlling for leverage. Once we control for idiosyncratic volatility, the return spread between high and low distress stocks reduces to -0.54%. Finally, controlling for profitability reduces the spread to -0.26% per month making it statistically insignificant. These results suggest that the return to high minus low distressed stock portfolios can be attributed to idiosyncratic volatility, leverage and profitability. The results are consistent with the notion that the distress risk anomaly is an amalgamation of other anomalies and return relationships previously documented in the literature.

3.4 Credit Spreads As a Measure of Default Risk

Given the results in the previous section, instead of using stock characteristics to measure financial distress, we take a different approach and use yields on corporate bonds in excess of the treasury rate to measure ex-ante probability of default. As mentioned earlier, this measure offers several advantages over others that have been used by previous papers. It is available in high frequency, which increases the power of statistical analyses we carry out. Unlike structural models of corporate bankruptcy that make simplifying assumptions about the capital structure of a firm, our proposed measure is model and assumption free. And unlike stock characteristics used to measure default risk, which may reflect information about future returns unrelated to distress risk, credit spreads reflect the market consensus view of the credit quality of the underlying firm.

There is now a significant body of theoretical research that shows that default-risk constitutes a considerable portion of credit spreads. Elton et al. (2001) report that default risk in credit spreads accounts for 19% to 41% of the spread level depending on company rating. Driessen (2005) also finds that default risk accounts for at the minimum18% (AA rated bonds) and as high as 52% (BBB rated bonds) of the corporate bond spread. Huang and Huang (2003) using the Longstaff-Schwartz model find that distress risk accounts for 39%, 34%, 41%, 73%, and 93% of the corporate bond spread respectively for bonds rated Aa, A, Baa, Ba and B. Longstaff, Mithal, and Neis (2005) use the information in credit default swaps (CDS) to obtain direct measures of the size of the default and non-default components in corporate spreads. They find that the default component represents 51% of the spread for AAA/AA rated bonds, 56% for A-rated bonds, 71% for BBB-rated bonds, and 83% for BB-rated bonds. The similarity in the information content of CDS spreads and bond credit spreads with respect to default is supported by Zhu (2005). He confirms, through co-integration tests, that the theoretical parity relationship between these two types of credit spreads holds as a long run equilibrium condition.[40]

3.4.1 Credit Spreads and Bankruptcy Prediction

Consistent with the studies discussed above, in this section we empirically show that bond spreads are a good ex-ante predictor of corporate defaults. In particular, we test to see if credit spreads improve default prediction beyond measures previously used in the

[40] In this study we have chosen to use bond spreads instead of CDS spreads because bond data is available for a substantially larger number of companies and is available for a much longer time period.

literature.[41] To measure the probability that a firm defaults, we estimate a dynamic panel model using a logit specification, following Shumway (2001), Chava and Jarrow (2004), CHS (2008) and others. We use information available at the end of the calendar year to predict defaults twelve months ahead. Specifically, the marginal probability of default (*PD*) for company *i* over the next year *t* is assumed to follow a logistic distribution:

$$PD_t^i = \frac{1}{1+\exp\left(-\alpha - \beta' X_t^i\right)} \tag{3.6}$$

where *X* is a vector of explanatory variables available at the time of prediction, and includes a comprehensive list of explanatory variables that have been used by previous papers to predict corporate bankruptcy. We use accounting variables used in calculating Altman's z-score, Ohlson's o-score, market based variables used by Shumway (2001) and CHS (2008), as well as Merton's distance-to-default measure. We also use Standard and Poor's (S&P) corporate ratings obtained from COMPUSTAT. All the variables used in the hazard regressions that follow are described in detail in the Appendix.

Table 3.5 reports results for the first set of hazard regressions. In the first column, we use the same covariates (*NIMTAVG, TLMTA, EXRETAVG, SIGMA, RSIZE, CASHMTA, MB* and *PRICE*) used in CHS (2008). The sample includes only firms that have issued bonds for the 1980 to 2008 time period. As a comparison, we report the estimates using the full sample (including firms that have not issued bonds), and also

---

[41] Bharath and Shumway (2008) document that credit spreads contain useful information in predicting defaults. In this paper, we significantly increase the number of defaults used in the hazard regressions, and also include a comprehensive list of alternative explanatory variables.

estimates from the CHS (2008) study in columns 7 and 6 respectively. The estimates from these three samples are very similar indicating that the bond dataset is not biased. When we use Merton's distance-to-default (*DD*) measure as a predictor, we obtain similar results to those in CHS (2008). Results from this regression are reported in column 4.

Next, we add corporate spreads (*SPREAD*) as an additional covariate to the CHS (2008) and the Merton specifications. The estimates from these two regressions are reported in columns 2 and 5 respectively. We also report estimates from a regression using *SPREAD* as the only covariate in column 3. Our proposed measure improves the explanatory power of both the CHS and Merton models. We report McFadden's pseudo R^2 coefficients for each regression.[42] The pseudo R^2 value increases from 27.6% for the CHS model to 37.4% for the CHS model used in conjunction with *SPREAD* in predicting bankruptcies. The specification that uses *SPREAD* alone has a pseudo R^2 coefficient of 26.5% which is similar to the pseudo R^2 for the CHS specification. Pseudo R^2 improves from 24.1% to 30.4% when Merton's *DD* is used in conjunction with *SPREAD*.

We also investigate whether it is appropriate to use corporate bond ratings as a measure of default risk. Many studies in this literature, including Avramov et al. (2006a), use corporate bond ratings as a proxy for distress risk. In this paper we show that *SPREAD* and *RATING* are not perfect substitutes. In fact, in Table 3.4 we show that there is much variation in credit spreads within a rating group. The correlation between credit spreads and ratings is only 0.45. AA- bonds, for instance, have an average credit spread of 84.30 basis points with a standard deviation of 43.93 basis points. A one

[42] McFadden's pseudo R^2 is calculated as $1 - L1/L0$, where L1 is the log likelihood of the estimated model and L0 is the log likelihood of a null model that includes only a constant term.

standard deviation move in credit spreads would firmly take an AA- bond's rating to a BBB+ rating which is 4 rating levels down. These results indicate that measuring default risk through company ratings can yield misleading results. This intuition is further supported by hazard regressions in columns 8 and 9 of Table 3.5. Pseudo R^2 improves from 23.6% to 30.5% when *RATING* is used in conjunction with *SPREAD*.

Table 3.6 further shows that adding *SPREAD* to Altman and Ohlson specifications have similar effects in improving the pseudo R^2 values. *SPREAD* enters with positive sign and has high statistical significance when used in conjunction with either of the models. Finally when we include all of the variables in Table 3.7, *SPREAD* enters with the expected sign and statistical significance while significantly improving the pseudo R^2. The analyses suggest that credit spread is an important predictor of corporate defaults and contains information related to financial distress not found in other measures commonly used in the literature.

3.4.2 Credit Spreads and Firm Characteristics

To see how corporate bond spreads are related to firm characteristics we form portfolios based on credit spreads. Each month from January 1981 through December 2008, companies in our sample are ranked and put into three portfolios based on the value of their credit spreads in the previous month. As described earlier, credit spreads are value-weighted averages of firms' outstanding bond spreads in a given month. For each portfolio, we calculate average book-to-market, size, momentum, and beta values for all the companies in that portfolio in a given month. Table 3.8 reports summary statistics for firm characteristics and value-weighted average monthly returns for credit spread portfolios. Credit spreads vary negatively with firm size and positively with book-to-market. The relationship with momentum is not monotonic, but the difference in past returns between the low and the high credit spread portfolios is positive and significant. In contrast to earlier studies, we find that equity returns increase monotonically with credit spreads.

3.4.3 Credit Spreads and Equity Returns

In this section we examine how corporate bond spreads are related to future realized equity returns. In particular we test whether stocks with high default risk as measured by credit spreads have anomalously low returns after controlling for standard risk factors. In the analyses that follow, we create two related but distinct proxies of credit risk. First, we use credit spreads, calculated as the difference between the corporate bond yield and the corresponding maturity matched treasury rate, to proxy for aggregate default risk. Second, we use credit spreads that are net of expected losses to proxy for each firm's exposure to the systematic component of default risk.

In order to calculate credit spreads that are net of expected losses we adopt a procedure used by Driessen et al. (2007), Elton et al. (2001) and Campello, Chen and Zhang (2004):

$$NetSpread_t = \left[PD \times \left(1 - L\right) + \left(1 - PD\right)\right] \times \left[1 + Spread_t\right] - 1 \qquad (3.7)$$

In Equation (7), *NetSpread* is the corporate bond spread net of expected losses, *PD* is the physical probability of default, *L* is the loss rate in the event of default, and *Spread* is the corporate bond credit spread calculated as the difference between the corporate bond yield and the corresponding maturity matched treasury rate. In Equation (7), we assume that default losses are incurred at maturity. We use CHS-score described in Section 3.2 to calculate physical probabilities of default. We follow Elton et al. (2001) and Driessen et al. (2007), and use historical loss rates reported in Altman and Kishmore (1998) by rating category. The loss rates vary from 32% for AAA-rated firms to 62% for CCC-rated firms.

We sort stocks into deciles each January from 1981 through 2008, according to the two distress measures calculated using corporate spreads. The stocks in each decile portfolio are held for a year. As before, if a delisting return is available we use the delisting return, otherwise we use the last available return in CRSP. To save space we only report returns for the top and bottom, and the difference between top and bottom deciles. The return results are reported in Table 3.9. The results under 'Bond Spreads' on the left hand side use credit spreads calculated as the difference between the corporate bond yield and the corresponding maturity matched treasury rate. The results under 'Bond Spreads In Excess of Expected Losses' on the right hand side use credit spreads that are net of expected losses.

Our results challenge those obtained in the previous studies. Using credit spreads, as a measure of default risk, the difference in raw returns between the highest and lowest default risk portfolios is 0.129% per month and statistically insignificant. The intercepts from the market and the 4-factor models are also economically and statistically insignificant. We find similar results when firms are sorted based on their exposures to the systematic component of default risk. The 4-factor monthly alphas for a portfolio formed by going long stocks in the highest distress portfolio and short stocks in the lowest distress risk portfolio are -0.208% and -0.156% using credit spreads and using credit spreads net of expected losses respectively.

There is a positive relationship between credit spreads and raw equity returns, but the return of the high minus low credit spread portfolio is not statistically significant. CAPM and the Carhart 4-factor regressions show that alphas are further subsumed in all credit spread portfolios suggesting that default risk is captured mainly by the market factor and partly by the size and the value factors. The size and value factors have statistically significant positive loadings for the highest credit risk portfolio, using either measure, suggesting that these factors are related to default risk. In 4-factor regressions the momentum factor has a negative and statistically significant loading in the highest credit risk portfolio regressions, consistent with the notion that poor performers of the past are likely to be today's distressed firms.

Ranking stocks on their real-world default probabilities, as done in Dichev (1998) and Campbell, Hilscher and Szilagyi (2008), implicitly assumes that high default probability stocks also have high exposure to the systematic component of default risk. Using corporate spreads we explicitly account for the systematic component in the risk of distress. To the best of our knowledge we are the first to explicitly rank equity returns according to firms' exposures to the systematic component of default risk. Overall, the results suggest that there is no evidence of default risk being negatively priced.

3.4.4 Robustness Checks

As we are using average spreads for each firm, to ensure that our results are not biased one way or another, in this section we consider the impact of bond liquidity and maturity on bond spreads and equity returns. We use some of the proxies utilized by Longstaff et al. (2005) in their study to measure corporate bond liquidity.[43] A dummy variable is given each month a value of one or zero depending on the characteristics of the underlying bond. We then add up the dummy variables to come up with an overall liquidity score. The first proxy is used to measure general availability of the bond issue in the market. If the outstanding market value of a bond is larger than the median market value of all bonds, then the dummy variable is assigned a value of one. The second proxy is the age of the bond and parallels the notion of on-the-run and off-the-run bonds in treasury markets, with on-the-run bonds being more liquid. If the age of a bond is less than the median age of all bonds, then the dummy variable is assigned a value of one. The third proxy is the time to maturity of the bond. It has been shown that there are maturity clienteles for corporate bonds and that shorter-maturity corporate bonds tend to be more liquid than longer-maturity bonds. If the time to maturity of a bond is less than seven years then the dummy variable is assigned a value of one. The fourth proxy that we use is a dummy variable for bonds rated by major rating agencies such as S&P and Moody's. If a bond is rated, then it is more likely to be liquid and the dummy variable is assigned a value of one. The maximum liquidity value assigned to a bond is four and the minimum liquidity value is zero.

[43] For a small subset of our sample, we have bid-ask, volume and turnover information. We carried out similar analyses described in this section and arrived at the same conclusions.

We divide our sample into three liquidity groups based on the liquidity score, and calculate average spread and one month ahead equity returns. The average spread for illiquid bonds is 50 basis points higher than for liquid bonds, and the difference is statistically significant. The differential for equity returns, on the other hand, is relatively small and insignificant. Portfolio returns are summarized in Table 3.10. In the monthly portfolios one can observe that the difference in raw returns between the highest and lowest default risk portfolios as well as the intercepts from the market and the 4-factor models for the high minus low credit risk portfolios are economically and statistically insignificant. This is true regardless of whether the underlying bonds are liquid or illiquid. These results indicate that liquidity effects are unlikely to be driving our findings.

As there are differences in values and variation in spreads across different bond maturities, in an effort to understand if the pricing of default risk varies across maturities, we split our sample into four maturity buckets: 1 to 4, 4 to 7, 7 to 11, and greater than 11 years. For each firm we calculate a weighted (by market value) average of bond spread within each time-to-maturity group. We carry out our analyses for each maturity bucket treating each company–maturity spread as a separate observation. We form three equally weighted portfolios of equity returns based on credit spread in each maturity group considered. Summary statistics of equity returns for company–maturity bucket / spread portfolios are reported in Table 3.11. In all time-to-maturity buckets, the difference in raw returns between the highest and lowest default risk portfolios as well as the intercepts from the market and the 4-factor models for the high minus low credit risk portfolios are economically and statistically insignificant. Since the uniform ranking of equity portfolio returns with respect to credit spreads yield similar patterns across different time-to-maturity groups, we conclude that our findings are not impacted by using an average credit spread.

3.5 Conclusion

In this paper we examine the pricing of default risk in equity returns. Our contribution to this literature is three-fold. First, we show that the distress risk anomaly is an amalgamation of other anomalies and return relationships previously documented in the literature. Second, ours is the first paper to use corporate bond spreads to measure the ex-ante probability of default risk. We show that in hazard rate regressions, credit spreads drive out the significance of most of the other measures that are used to predict corporate defaults and significantly improve the pseudo R^2 values in all specifications. Third, contrary to previous findings, we show that default risk is not priced negatively in the cross section of equity returns. We sort firms according to their exposures to the systematic component of default risk as well as their aggregate default risk. To the best of our knowledge we are the first to explicitly rank equity returns according to firms' exposures to the systematic component of default risk. Portfolios sorted both on credit spreads and on credit spreads net of expected losses have positive raw returns but do not deliver significant positive or negative returns after controlling for well known risk factors. Our findings challenge the previous studies that have found an anomalous relationship between credit risk and equity returns. The analyses in this paper take the right step towards finding a more appropriate measure of systematic default risk that can explain the cross section of equity returns in line with the rational expectations theory.

Table 3.1: Summary Statistics

Table 3.1 reports summary statistics for firm characteristics and distress measures for companies in the CRSP sample (left panel) and the Bond sample (right panel). *tlmta* is the ratio of total liabilities to the market value of total assets, *cashmta* is the ratio of cash to the market value of total assets, *exretavg* is a geometrically declining average of monthly log excess stock returns relative to the S&P 500 index, *nimtaavg* is a geometrically declining average of past values of the ratio of net income to the market value of total assets, *rsize* is the log ratio of market capitalization to the market value of the S&P 500 index, *totvol* is the standard deviation of daily stock returns over the previous calendar year, idiovol is the standard deviation of regression errors obtained from regressing daily excess returns on the Fama and French (1993) factors, *price* is the log price per share truncated from above at $15, *CHS-PD*100* is the CHS probability of default reported as a percentage. *Merton-DD* is the Merton distance-to-default measure. ta/cpi is total assets divided by consumer price index, *tl/ta* is the ratio of total assets to total liabilities, *cl/ca* is the ratio of current liabilities to current assets, *wc/ta* is the ratio of working capital to total assets, *ni/ta* is the ratio of net income to total assets, *ffops/tl* is the ratio of funds from operations to total liabilities, *Δni* is the change in net income, *rearn/ta* is the ratio of retained earnings to total assets, *ebit/ta* is the ratio of earnings before interest and taxes to total assets, *me/tl* is the ratio of market equity to total liabilities, and *sales/ta* is the ratio of sales to total assets. How these variables calculated are described in the appendix. P25, P50 and P75 represent 25[th], 50[th] and 75[th] percentiles respectively.

	CRSP SAMPLE						BOND SAMPLE				
Variable	Mean	Std Dev	P25	P50	P75	Variable	Mean	Std Dev	P25	P50	P75
	CHS Variables and Stock Characteristics						*CHS Variables and Stock Characteristics*				
mb	1.95	1.43	0.91	1.52	2.57	mb	1.76	1.14	0.97	1.45	2.23
me	1156.48	7848.51	24.67	93.11	405.72	me	5104.56	17098.45	322.17	1115.78	3524.17
cashmta	0.089	0.092	0.021	0.070	0.113	cashmta	0.051	0.060	0.010	0.028	0.070
exretavg	-0.009	0.043	-0.032	-0.005	0.018	exretavg	-0.002	0.032	-0.018	0.000	0.016
nimtavg	0.003	0.014	0.000	0.005	0.012	nimtavg	0.007	0.009	0.003	0.008	0.012
rsize	-10.44	1.76	-11.77	-10.56	-9.22	rsize	-8.19	1.31	-8.95	-7.85	-7.16
idiovol	0.034	0.027	0.017	0.027	0.042	idiovol	0.019	0.013	0.012	0.016	0.022
totvol	0.036	0.027	0.019	0.029	0.045	totvol	0.022	0.014	0.014	0.018	0.025
price	2.166	0.694	1.749	2.546	2.708	price	2.592	0.350	2.708	2.708	2.708
CHS PD * 100	0.081	0.163	0.021	0.038	0.075	CHS PD * 100	0.052	0.104	0.020	0.032	0.051
	Merton Model Variables						*Merton Model Variables*				
Merton DD	7.267	36.504	3.108	5.358	8.626	Merton DD	8.113	5.785	4.776	7.309	10.411
	Ohlson Model Variables						*Ohlson Model Variables*				
ta / cpi	3.947	2.162	2.385	3.801	5.376	ta / cpi	6.588	1.592	5.606	6.605	7.624
tl / ta	0.275	0.147	0.154	0.288	0.400	tl / ta	0.336	0.107	0.264	0.352	0.419
wcap / ta	0.142	0.414	0.018	0.126	0.260	wcap / ta	0.021	0.166	-0.086	0.020	0.117
cl / ca	1.247	61.853	0.305	0.492	0.751	cl / ca	3.656	71.806	0.505	0.707	0.991
ni / ta	0.001	0.064	0.000	0.004	0.009	ni / ta	0.005	0.012	0.002	0.005	0.009

ffops / tl	0.070	6.193	-0.002	0.022	0.070
Δ ni	0.025	0.526	-0.190	0.023	0.268
Altman Model Variables					
wcap / ta	0.142	0.414	0.018	0.126	0.260
rearn / ta	0.016	3.887	-0.014	0.058	0.158
ebit / ta	0.007	0.030	0.002	0.010	0.017
me / tl	11.971	273.842	0.526	1.495	4.624
sales / ta	0.075	0.217	0.009	0.031	0.084

ffops / tl	0.037	0.073	0.006	0.025	0.051
Δ ni	0.026	0.464	-0.120	0.035	0.209
Altman Model Variables					
wcap / ta	0.021	0.166	-0.086	0.020	0.117
rearn / ta	0.094	0.139	0.030	0.085	0.159
ebit / ta	0.013	0.010	0.007	0.012	0.017
me / tl	1.686	7.843	0.389	0.825	1.698
sales / ta	0.013	0.026	0.002	0.005	0.013

Table 3.2: Distress Portfolio Returns

Table 3.2 reports CAPM and 4-factor regression results for distress portfolios. We sort stocks into deciles each January from 1981 through December 2008, according to their default probabilities calculated using the CHS hazard model, Ohlson's o-score, and Merton's distance-to-default measure. We compute the value-weighted return for these decile portfolios on a monthly basis and regress the portfolio return in excess of risk-free rate on the market (*MKT*), size (*SMB*), value (*HML*), and momentum (*MOM*) factors. The factors are obtained from Ken French's website. The results under 'Bond Sample' on the right hand side include only the companies in our bond sample. We report regression results for only the top and bottom decile portfolios to save space. Panels A, B and C report regression results for portfolios constructed using the CHS hazard model, Ohlson's o-score, and Merton's distance-to-default measure respectively. Absolute values of *t-statistics* are reported in parentheses below coefficient estimates. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: CHS Probability of Default

CRSP Sample

	Alpha * 100	MKT	SMB	HML	MOM
10th	0.598				
	(1.96)**				
	0.151	1.051			
	(0.90)	(28.09)***			
	0.078	0.962	0.104	-0.347	0.326
	(0.56)	(28.68)***	(2.35)**	(6.73)***	(10.34)***
	Alpha * 100	MKT	SMB	HML	MOM
90th	-0.626				
	(1.31)				
	-1.255	1.480			
	(3.96)***	(20.92)***			
	-0.752	1.335	0.923	0.240	-0.810
	(3.28)***	(23.92)***	(12.51)***	(2.80)***	(15.42)***
	Alpha * 100	MKT	SMB	HML	MOM
90th - 10th	-1.224				
	(2.98)***				
	-1.406	0.428			
	(3.52)***	(4.80)***			
	-0.830	(0.37)	0.819	0.587	-1.136
	(2.94)***	(5.44)***	(9.02)***	(5.56)***	(17.58)***

Bond Sample

	Alpha * 100	MKT	SMB	HML	MOM
10th	0.455				
	(2.03)**				
	0.111	0.807			
	(1.04)	(33.69)***			
	0.019	0.900	-0.269	0.172	0.001
	(1.20)	(39.28)***	(8.89)***	(4.89)***	(0.05)
	Alpha * 100	MKT	SMB	HML	MOM
90th	0.270				
	(1.72)*				
	-0.514	1.239			
	(1.96)**	(25.15)***			
	-0.302	1.383	0.099	0.668	-0.483
	(2.55)***	(35.27)***	(1.90)**	(11.08)***	(13.10)***
	Alpha * 100	MKT	SMB	HML	MOM
90th --10th	-0.185				
	(1.66)*				
	-0.625	0.432			
	(1.93)**	(6.43)***			
	-0.321	0.483	0.368	0.496	-0.484
	(2.15)**	(10.05)***	(6.02)***	(9.51)***	(14.66)***

Table 3.2: Distress Portfolio Returns

Table 3.2 reports CAPM and 4-factor regression results for distress portfolios. We sort stocks into deciles each January from 1981 through December 2008, according to their default probabilities calculated using the CHS hazard model, Ohlson's o-score, and Merton's distance-to-default measure. We compute the value-weighted return for these decile portfolios on a monthly basis and regress the portfolio return in excess of risk-free rate on the market (*MKT*), size (*SMB*), value (*HML*), and momentum (*MOM*) factors. The factors are obtained from Ken French's website. The results under 'Bond Sample' on the right hand side include only the companies in our bond sample. We report regression results for only the top and bottom decile portfolios to save space. Panels A, B and C report regression results for portfolios constructed using the CHS hazard model, Ohlson's o-score, and Merton's distance-to-default measure respectively. Absolute values of *t-statistics* are reported in parentheses below coefficient estimates. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: CHS Probability of Default

CRSP Sample

	Alpha * 100	MKT	SMB	HML	MOM
10th	0.598				
	(1.96)**				
	0.151	1.051			
	(0.90)	(28.09)***			
	0.078	0.962	0.104	-0.347	0.326
	(0.56)	(28.68)***	(2.35)**	(6.73)***	(10.34)***
	Alpha * 100	MKT	SMB	HML	MOM
90th	-0.626				
	(1.31)				
	-1.255	1.480			
	(3.96)***	(20.92)***			
	-0.752	1.335	0.923	0.240	-0.810
	(3.28)***	(23.92)***	(12.51)***	(2.80)***	(15.42)***
	Alpha * 100	MKT	SMB	HML	MOM
90th - 10th	-1.224				
	(2.98)***				
	-1.406	0.428			
	(3.52)***	(4.80)***			
	-0.830	(0.37)	0.819	0.587	-1.136
	(2.94)***	(5.44)***	(9.02)***	(5.56)***	(17.58)***

Bond Sample

	Alpha * 100	MKT	SMB	HML	MOM
10th	0.455				
	(2.03)**				
	0.111	0.807			
	(1.04)	(33.69)***			
	0.019	0.900	-0.269	0.172	0.001
	(1.20)	(39.28)***	(8.89)***	(4.89)***	(0.05)
	Alpha * 100	MKT	SMB	HML	MOM
90th	0.270				
	(1.72)*				
	-0.514	1.239			
	(1.96)**	(25.15)***			
	-0.302	1.383	0.099	0.668	-0.483
	(2.55)***	(35.27)***	(1.90)**	(11.08)***	(13.10)***
	Alpha * 100	MKT	SMB	HML	MOM
90th --10th	-0.185				
	(1.66)*				
	-0.625	0.432			
	(1.93)**	(6.43)***			
	-0.321	0.483	0.368	0.496	-0.484
	(2.15)**	(10.05)***	(6.02)***	(9.51)***	(14.66)***

Panel B: Merton Distance to Default

	CRSP Sample					Bond Sample					
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
10th	-0.175					10th	0.521				
	(0.39)						(2.38)**				
	-0.796	1.442					-0.124	1.375			
	(2.80)***	(22.99)***					(1.70)*	(35.13)***			
	-0.450	1.259	1.059	0.034	-0.495		-0.002	1.453	0.121	0.334	-0.176
	(2.07)**	(24.27)***	(15.08)***	(0.43)	(10.16)***		(1.88)*	(35.02)***	(2.18)**	(5.22)***	(4.57)***
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
90th	0.579					90th	0.482				
	(2.67)***						(2.36)**				
	0.166	0.842					0.155	0.699			
	(2.35)**	(53.90)***					(1.34)	(27.39)***			
	0.165	0.877	-0.202	0.008	0.015		0.097	0.768	-0.309	0.063	0.043
	(2.50)**	(55.68)***	(9.45)***	(0.31)	(1.02)		(1.91)*	(29.90)***	(8.96)***	(1.58)	(1.80)*
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
90th - 10th	0.753					90th - 10th	-0.039				
	(1.94)*						(1.14)				
	0.962	-0.599					0.279	-0.676			
	(4.05)***	(19.68)***					(2.21)**	(13.28)***			
	0.615	-0.381	-1.261	-0.027	0.510		0.099	-0.685	-0.431	-0.271	0.219
	(3.10)***	(17.21)***	(8.27)***	(0.93)	(6.42)***		(1.80)*	(12.69)***	(5.94)***	(3.26)***	(4.36)***

Panel C: Ohlson O-Score

	CRSP Sample						Bond Sample				
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
10th	0.637					10th	0.459				
	(1.95)*						(2.02)**				
	0.036	1.175					0.067	0.837			
	(0.23)	(34.43)***					(0.65)	(36.69)***			
	0.417	1.025	-0.055	-0.506	-0.068		0.140	0.844	-0.345	-0.150	0.026
	(2.80)***	(28.72)***	(1.15)	(9.20)***	(2.05)**		(1.51)	(38.18)***	(11.62)***	(4.40)***	(1.26)
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
90th	-0.643					90th	0.268				
	(1.37)						(1.91)*				
	-1.119	1.252					0.153	0.885			
	(3.13)***	(15.84)***					(0.74)	(19.32)***			
	-0.898	0.957	1.284	-0.209	-0.304		-0.103	1.063	0.260	0.713	-0.155
	(3.01)***	(13.39)***	(13.36)***	(1.89)*	(4.57)***		(0.56)	(24.10)***	(4.39)***	(10.50)***	(3.78)***
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
90th - 10th	-1.280					90th - 10th	-0.191				
	(2.28)**						(1.44)				
	-1.155	0.077					0.086	0.048			
	(2.60)***	(8.85)***					(0.35)	(0.88)			
	-1.280	-0.068	1.339	0.296	-0.236		-0.243	0.219	0.605	0.863	-0.181
	(3.43)***	(5.48)***	(2.91)***	(6.62)***	(2.39)**		(1.15)	(4.35)***	(8.95)***	(11.13)***	(3.86)***

Table 3.3: Stock Characteristics and Default Risk

Table 3.3 shows the 4-factor alphas for distress portfolios before and after controlling for idiosyncratic volatility, profitability and leverage. Distress portfolios are formed by sorting stocks into five groups each January from 1981 to 2008 according to the CHS probability of default. Then within each default group we first sort stocks based on the previous year's idiosyncratic volatility into five groups creating a total of 25 portfolios. The five distress portfolios are averaged over each of the idiosyncratic volatility portfolios to account for the impact of idiosyncratic volatility. Finally we calculate the 4-factor alphas for the distress portfolios as well as the high distress-low distress hedge portfolio. The same procedure is repeated for profitability and leverage characteristics and we report only the 4-factor alphas for distress portfolios as well as hedge portfolios that have been controlled for the effects of the aforementioned stock characteristics. Idiosyncratic volatility is calculated relative to the Fama-French 3-factor model as in AHXZ (2006). Profitability is measured using *NIMTAVG*, and leverage is measured using *TLMTA*. *NIMTAAVG* is a geometrically declining average of past values of the ratio of net income to the market value of total assets, and *TLMTA* is the ratio of total liabilities to the market value of total assets. Absolute values of *t-statistics* are reported in parentheses below coefficient estimates. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Panel A: 4-Factor Returns						
4-Factor Alphas (*100) Before/After Controlling for Stock Characteristics						
	L	**2**	**3**	**4**	**H**	**H-L**
Before controls	0.079	0.133	0.014	-0.158	-0.803	-0.882
	-0.71	(1.94)*	-0.13	-1.03	(3.29)***	(2.71)***
Controlling for Idio Volatility	-0.091	-0.219	-0.304	-0.279	-0.627	-0.537
	-0.62	(1.88)*	(2.73)***	(2.01)**	(3.17)***	(2.08)**
Controlling for Profitability	0.012	-0.104	-0.006	0.008	-0.251	-0.263
	(0.14)	(1.89)*	(0.08)	(0.08)	(1.74)*	(1.39)
Controlling for Leverage	0.072	-0.006	0.004	-0.122	-0.545	-0.617
	(0.98)	(0.1)*	(0.05)	(1.1)	(3.01)***	(2.93)***

Panel B: Stock Characteristics						
Idiosyncratic Volatility	0.025	0.026	0.027	0.032	0.045	0.019
Profitability	0.012	0.008	0.005	0.001	-0.011	-0.022
Leverage	0.216	0.333	0.456	0.550	0.605	0.389

Table 3.4: Credit spread by rating categories

Table 3.4 reports summary statistics for credit spreads by rating category. The benchmark risk-free yield is the yield of the closest maturity treasury. We include only straight fixed-coupon corporate bonds for the January 1974-December 2008 time period. Bonds for financial firms are excluded. The spreads are given in annualized basis points and ratings in this sample come from Standard and Poor's.

Rating Category (S&P)	Number of Observations	Mean Spread (bps)	Std Dev Spread (bps)
AAA	1157	64.30	27.47
AA+	316	87.58	32.07
AA	2973	77.51	35.70
AA-	2966	84.30	43.93
A+	5155	96.99	45.77
A	7778	102.28	51.99
A-	5397	112.24	61.65
BBB+	4801	124.45	67.24
BBB	4882	146.47	88.86
BBB-	3559	185.86	113.99
BB+	1224	272.54	142.87
BB	949	321.31	134.27
BB-	709	384.52	142.45
B+	342	405.91	129.51
B	266	448.77	156.50
B-	57	508.09	148.10
CCC+	34	455.60	117.19
CCC	29	583.79	116.17
All Ratings	42605	133.67	104.39

Table 3.5: Bankruptcy Prediction – CHS Covariates, Ratings and Distance-to-Default

Table 3.5 reports results from logit regressions of the bankruptcy indicator on predictor variables. *NIMTAAVG* is a geometrically declining average of past values of the ratio of net income to the market value of total assets, *TLMTA* is the ratio of total liabilities to the market value of total assets, *EXRETAVG* is a geometrically declining average of monthly log excess stock returns relative to the S&P 500 index, *SIGMA* is the standard deviation of daily stock returns over the previous three months, *RSIZE* is the log ratio of market capitalization to the market value of the S&P 500 index, *CASHMTA* is the ratio of cash to the market value of total assets, *MB* is the market-to-book ratio, *PRICE* is the log price per share truncated from above at $15. These variables are described in detail in the appendix. *Rating* is the Standard and Poor's (S&P) corporate rating obtained from COMPUSTAT. Results under 'All Firms' are estimates computed using the full sample of defaults with available accounting information. Results under 'Firms with bonds' are estimates computed using the sample of defaults from companies that have issued bonds with available accounting information. Results under 'CHS sample' shows the estimates CHS report in their paper. Absolute values of z-statistics are reported in parentheses below coefficient estimates. McFadden pseudo R^2 values are reported for each regression. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	(8)	(9)	(10)	(10)	(11)	(12)
Sample period:	1981-2008	1981-2008	1981-2008	1981-2008	1981-2008	1981-2008
NIMTAAVG					-15.667	-12.039
					(1.28)	(1.40)
TLMTA					1.890	1.205
					(1.60)	(2.34)**
EXRETAVG					-15.753	-16.015
					(4.31)***	(5.34)***
SIGMA					0.692	0.037
					(0.84)	(0.43)
RSIZE					-0.233	-0.330
					(1.09)	(1.09)
CASHMTA					-2.064	-2.657
					(1.11)	(1.11)
MB					-0.009	0.055
					(0.27)	(0.27)
PRICE					0.022	0.188
					(0.31)	(0.31)
SPREAD		17.870		15.229		14.600
		(6.43)***		(4.34)***		(3.19)***
DD			-0.666	-0.556	-0.260	-0.302

			(5.70)***	(6.14)***	(1.74)*	(1.78)*
RATING	0.410	0.257	0.122	0.015	0.086	-0.014
	(13.26)***	(6.98)***	(2.47)**	(0.30)	(1.12)	(0.15)
CONSTANT	-9.149	-8.116	-3.154	-3.017	-8.464	-8.286
	(21.69)***	(18.90)***	(3.78)***	(4.21)***	(3.07)***	(2.74)***
Observations	8068	8068	6814	6814	6736	6736
Bankruptcies	77	77	51	51	51	51
Pseudo R^2	0.236	0.305	0.279	0.315	0.351	0.377
Sample Type	Firms with Bonds	Firms with Bonds	Firms with Bonds	Firms with Bonds	Firms with Bonds	Firms with Bonds

Table 3.5 continued: Bankruptcy Prediction – Ratings, Spreads and Distance-to-Default

Table 3.5 reports results from logit regressions of the bankruptcy indicator on predictor variables. *NIMTAAVG* is a geometrically declining average of past values of the ratio of net income to the market value of total assets, *TLMTA* is the ratio of total liabilities to the market value of total assets, *EXRETAVG* is a geometrically declining average of monthly log excess stock returns relative to the S&P 500 index, *SIGMA* is the standard deviation of daily stock returns over the previous three months, *RSIZE* is the log ratio of market capitalization to the market value of the S&P 500 index, *CASHMTA* is the ratio of cash to the market value of total assets, *MB* is the market-to-book ratio, *PRICE* is the log price per share truncated from above at $15. These variables are described in detail in the appendix. *Rating* is the Standard and Poor's (S&P) corporate rating obtained from COMPUSTAT. Results under 'All Firms' are estimates computed using the full sample of defaults with available accounting information. Results under 'Firms with bonds' are estimates computed using the sample of defaults from companies that have issued bonds with available accounting information. . Results under 'CHS sample' shows the estimates CHS report in their paper. Absolute values of z-statistics are reported in parentheses below coefficient estimates. McFadden pseudo R^2 values are reported for each regression. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	(8)	(9)	(10)	(10)	(11)	(12)
Sample period:	1981-2008	1981-2008	1981-2008	1981-2008	1981-2008	1981-2008
NIMTAAVG					-15.667	-12.039
					(1.28)	(1.40)
TLMTA					1.890	1.205
					(1.60)	(2.34)**
EXRETAVG					-15.753	-16.015
					(4.31)***	(5.34)***
SIGMA					0.692	0.037
					(0.84)	(0.43)
RSIZE					-0.233	-0.330
					(1.09)	(1.09)
CASHMTA					-2.064	-2.657
					(1.11)	(1.11)
MB					-0.009	0.055
					(0.27)	(0.27)
PRICE					0.022	0.188
					(0.31)	(0.31)
SPREAD		17.870		15.229		14.600
		(6.43)***		(4.34)***		(3.19)***
DD			-0.666	-0.556	-0.260	-0.302
			(5.70)***	(6.14)***	(1.74)*	(1.78)*
RATING	0.410	0.257	0.122	0.015	0.086	-0.014
	(13.26)***	(6.98)***	(2.47)**	(0.30)	(1.12)	(0.15)
CONSTANT	-9.149	-8.116	-3.154	-3.017	-8.464	-8.286
	(21.69)***	(18.90)***	(3.78)***	(4.21)***	(3.07)***	(2.74)***

Observations	8068	8068	6814	6814	6736	6736
Bankruptcies	77	77	51	51	51	51
Pseudo R^2	0.236	0.305	0.279	0.315	0.351	0.377
Sample Type	Firms with Bonds	Firms with Bonds	Firms with Bonds	Firms with Bonds	Firms with Bonds	Firms with Bonds

Table 3.6: Bankruptcy Prediction – Altman and Ohlson Covariates

Table 3.6 reports results from logit regressions of the bankruptcy indicator on predictor variables. *WCTA* is the ratio of working capital to total assets, *RETA* is the ratio of retained earnings to total assets, *EBITTA* is the ratio of earnings before interest and taxes to total assets, *METL* is the ratio of market equity to total liabilities, *STA* is the ratio of sales to total assets, *SIZE* is total assets divided by the consumer price index, *CLCA* is the ratio of current liabilities to current assets, *OENEG* is a dummy variable equal to one if total liabilities exceeds total assets and zero otherwise, *NITA* is the ratio of net income to total assets, *FUTL* is the ratio of funds from operations to total liabilities, *INTWO* is a dummy variable equal to one if net income was negative for the past two years and zero otherwise, and *CHIN* is a measure of the change in net income. These variables are described in detail in the appendix. Absolute values of z-statistics are reported in parentheses below coefficient estimates. McFadden pseudo R^2 values are reported for each regression. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)
Sample period:	1981-2008	1981-2008	1981-2008	1981-2008
SIZE	-0.254	-0.208		
	(2.38)**	(1.67)*		
TLTA	20.372	14.304		
	(4.80)***	(3.54)***		
WCTA	0.068	-0.348		
	(0.09)	(0.63)		
CLCA	-0.002	-0.112		
	(1.88)*	(0.51)		
NITA	6.441	7.126		
	(0.35)	(0.35)		
FUTL	-8.076	-8.044		
	(1.15)	(1.07)		
CHIN	-0.300	-0.355		
	(1.31)	(1.37)		
INCDUM	0.905	0.600		
	(2.76)***	(1.65)*		
TEDUM	1.095	0.904		
	(2.69)**	(1.83)*		

WCTA			0.815	0.203
			(0.77)	(0.24)
RETA			-2.453	-0.530
			(2.28)**	(0.44)
EBITA			-24.779	-22.096
			(1.78)*	(1.61)
METL			-2.947	-1.737
			(3.31)***	(2.52)**
STA			28.703	30.320
			(1.32)	(1.46)
SPREAD		15.011		20.168
		(4.02)***		(5.20)***
CONSTANT	-11.409	-9.640	-2.977	-4.291
	(6.70)***	(6.29)***	(9.65)***	(8.87)***
Observations	6349	6349	5896	5896
Bankruptcies	51	51	48	48
Pseudo R^2	0.245	0.324	0.179	0.277
Sample Type	Firms with Bonds	Firms with Bonds	Firms with Bonds	Firms with Bonds

Table 3.7: Bankruptcy Prediction – All Covariates

Table 3.7 reports results from logit regressions of the bankruptcy indicator on predictor variables. The explanatory variables are all the covariates described in Tables 5 and 6. Absolute values of z-statistics are reported in parentheses next to coefficient estimates. McFadden pseudo R^2 values are reported for each regression. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

	(1)		(2)	
Sample period:	1981-2008		1981-2008	
NIMTAAVG	31.04	(1.48)	44.82	(1.89)*
TLMTA	1.39	(0.12)	4.89	(0.38)
EXRETAVG	-12.93	(2.81)***	-13.98	(2.90)***
SIGMA	-0.05	(0.04)	-1.08	(0.79)
RSIZE	-0.89	(2.47)**	-1.15	(3.09)***
CASHMTA	-6.09	(1.40)	-8.31	(1.43)
MB	-0.44	(2.28)**	-0.47	(2.31)**
PRICE	-0.06	(0.12)	0.07	(0.12)
DD	-0.31	(1.49)	-0.37	(1.52)
RATING	0.09	(0.86)	-0.04	(0.33)
SIZE	0.82	(2.44)**	1.00	(3.03)***
TLTA	-10.48	(0.29)	-30.15	(0.71)
WCTA	0.29	(0.30)	-0.17	(0.17)
CLCA	0.14	(0.65)	-0.09	(0.29)
NITA	-14.29	(1.19)	-19.27	(1.35)
FUTL	-2.35	(0.50)	-1.84	(0.32)
CHIN	-0.42	(1.66)*	-0.37	(1.38)
INCDUM	0.82	(1.77)*	0.77	(1.52)
TEDUM	2.55	(3.28)***	3.05	(3.45)***
RETA	1.75	(1.06)	1.53	(0.42)
EBITA	-1.99	(0.11)	-10.74	(0.57)
STA	-0.37	(0.35)	-1.38	(0.89)
METL	40.10	(1.55)	48.21	(1.68)*
SPREAD			17.97	(3.59)***
CONSTANT	-14.53	(0.66)	-10.57	(1.11)
Observations	5175		5175	
Bankruptcies	43		43	
Pseudo R^2	.415		.455	
Sample Type	Firms with Bonds		Firms with Bonds	

Table 3.8: Firm characteristics in credit-spread portfolios

In Table 3.8 we report firm characteristics such as that month's equity return, market capitalization (in $millions), book to market value, momentum and firm beta for three credit-spread portfolios. Each month from January 1981 through December 2008, value-weighted credit spread portfolios are formed from all stocks with available bond data using CRSP returns. Firms must also have Compustat data to calculate book-to-market values. Size is the market value of equity in millions of dollars and is taken from CRSP as the product of share price at the end of the month and the number of shares outstanding. Book-to-market (*BM*) is calculated as the ratio of book equity in the previous calendar month to market equity in the previous month for all stocks with Compustat data as well as credit spread information. Book equity value used in that month must have been available to the public for a minimum of 6 months. Previous return is the compounded raw returns of the past 12 months. We calculate each firm's beta for month t by regressing each stock's monthly returns on the value-weighted NYSE/AMEX index during the past 36 months. Beta is the regression coefficient on the value-weighted NYSE/AMEX index.

Spread Rank	Variable	Mean	Std Dev
Low	Return	0.00986	0.0655
	Size	26,237	64,575
	BM	0.48695	0.30274
	Prev Return	0.17002	0.24911
	Beta	0.93860	0.48353
Intermediate	Return	0.01307	0.07279
	Size	14,130	46449
	BM	0.61622	0.42316
	Prev Return	0.17671	0.27025
	Beta	0.98480	0.49288
High	Return	0.01359	0.10542
	Size	5,927	21647
	BM	0.83271	0.64552
	Prev Return	0.15031	0.40985
	Beta	1.09971	0.64248

Table 3.9: Monthly equity returns for credit spread portfolios

In Table 3.9 we report CAPM and 4-factor regression results for distress portfolios. We sort stocks into deciles each January from 1981 through December 2008, according to their credit spreads obtained at the beginning of December of the most recent year ended. We compute the value-weighted return for these decile portfolios on a monthly basis and regress the portfolio return in excess of risk-free rate on the market (*MKT*), size (*SMB*), value (*HML*), and momentum (*MOM*) factors. The factors are obtained from Ken French's website. The results under 'Bond Spreads' on the left hand side use credit spreads calculated as the difference between the corporate bond yield and the corresponding maturity matched treasury rate. The results under 'Bond Spreads In Excess of Expected Losses' on the right hand side use credit spreads that are net of expected losses. The 'Bond Spread' variable is a measure of the total default risk while the 'Bond Spreads In Excess of Expected Losses' proxy for only the systematic portion of default risk. We report regression results for only the top and bottom decile portfolios to save space. Absolute values of *t-statistics* are reported in parentheses below coefficient estimates. Statistical significance at the 10%, 5% and 1% levelsis denoted by *, **, and ***, respectively.

Monthly Equity Returns For Default Risk Portfolios

	Bond Spreads						**Bond Spreads In Excess of Expected Losses**				
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
10th	0.094					10th	0.090				
	(1.43)						(1.36)				
	-0.316	0.843					-0.321	0.841			
	(0.16)	(22.64)***					(0.08)	(22.21)***			
	-0.421	0.882	-0.35	-0.04	-0.02		-0.420	0.886	-0.342	-0.018	-0.023
	(0.38)	(22.61)***	(6.78)**	(0.66)	(0.44)		(0.25)	(22.22)***	(6.59)***	(0.29)	(0.62)
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
90th	0.223					90th	0.311				
	(0.49)						(0.69)				
	-0.233	1.063					-0.140	1.055			
	(0.65)	(13.39)***					(0.39)	(13.18)***			
	-0.629	1.272	0.432	0.976	-0.148		-0.576	1.278	0.416	1.008	-0.124
	(1.88)*	(15.68)***	(4.10)***	(7.89)***	(1.96)**		(1.72)*	(15.79)***	(3.95)***	(8.16)***	(1.65)*
	Alpha * 100	MKT	SMB	HML	MOM		Alpha * 100	MKT	SMB	HML	MOM
90th - 10th	0.129					90th - 10th	0.221				
	(0.05)						(0.05)				
	0.083	0.477					-0.181	0.048			
	(0.70)	(8.85)***					(0.69)	(6.67)***			
	-0.208	0.516	-0.063	0.109	-0.025		-0.156	0.219	0.605	0.863	-0.181
	(0.79)	(6.56)***	(0.62)	(0.91)	(0.35)		(0.84)	(6.50)***	(0.70)	(1.03)	(0.10)

Table 3.10: Monthly equity returns for bond liquidity / credit spread portfolios

In Table 3.10, we report one month ahead equity returns of credit-spread sorted portfolios for companies associated with different levels of bond market liquidity. We separately report equity returns for companies that are associated with high liquidity in the bond market as well as for companies that are associated with low liquidity in the bond market. In order to determine a bond's market liquidity level we use 4 proxies as described in the text. A dummy variable is given each month a value of one or zero depending on the characteristics of the underlying bond. We then add up the dummy variables to come up with an overall liquidity score for each bond. We value weight the liquidity scores of the bonds that belong to the same firm and assign each firm a single bond market liquidity measure in a given month. Weights are the outstanding market values of the bonds. In a similar fashion we calculate firm level credit spreads for each firm on a monthly basis. Every month, we group firms into three buckets based on their bond market liquidity level. Furthermore, within each bond market liquidity bucket, firms are grouped in to three portfolios based on their value weighted credit spreads. For each liquidity bucket we report uniformly ranked monthly raw returns for the three credit-spread portfolios, as well as raw return differences, CAPM and 4-factor Carhart model based monthly alphas between high credit spread and low credit spread portfolios. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Bond Liquidity Rank	Spread Rank	Avg Return	t-stat
High	L	0.8600	3.29***
	2	1.1900	3.82***
	H	1.0700	2.86***
	Raw Alpha H-L	0.0500	0.22
	CAPM Alpha H-L	-0.0810	-0.34
	Carhart Alpha H-L	-0.0290	0.14
Intermediate	L	0.7000	2.51**
	2	0.5400	1.78*
	H	0.9800	2.51**
	Raw Alpha H-L	0.1388	0.54
	CAPM Alpha H-L	0.0200	0.08
	Carhart Alpha H-L	0.0165	0.069
Low	L	1.0537	4.10***
	2	1.0570	3.73***
	H	0.9353	2.49**
	Raw Alpha H-L	-0.1184	-0.49
	CAPM Alpha H-L	-0.2260	-0.96
	Carhart Alpha H-L	-0.3190	-1.49

Table 3.11: Monthly equity returns for credit spread/maturity portfolios

In Table 3.11, we report returns of credit-spread sorted portfolios in different time-to-maturity groups. Maturity is the remaining time to maturity in years of the bonds. We allocate each bond to one of four maturity groups: Bucket 1 includes bonds with maturities less than 4 years but more than 1 year, Bucket 2 includes bonds with maturities greater than 4 years but less than 7 years, Bucket 3 includes bonds with maturities greater than 7 years but less than 11 years, and Bucket 4 includes bonds with maturities greater than 11 years. Each month from January 1981 through December 2008 bonds are assigned to four groups based on their time to maturity. For each firm we calculate four different credit-spread values: one for each maturity bucket. All credit spreads are value-weighted with respect to the market values of a firm's outstanding bonds. If a firm doesn't have any bonds outstanding in a given maturity bucket then it is excluded from the analysis regarding that time to maturity group. Within each maturity bucket firms are assigned to three portfolios based on their credit spreads. For each time-to-maturity bucket we calculate equal-weighted subsequent realized monthly equity returns for each credit-spread portfolio. In each maturity bucket we ask whether portfolios with high credit spread have unusually high or low returns relative to the predictions of standard asset pricing models such as the CAPM, and the four-factor Carhart model. We report uniformly ranked monthly raw returns for the three credit-risk portfolios, as well as raw return differences, CAPM and 4-factor Carhart model based monthly alphas between high credit spread and low credit spread portfolios. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Maturity Groups	spread -rank	Mean	t-value
Maturity-Bucket 1	L	1.0268	3.45***
	2	1.2140	3.56***
	H	1.0767	2.70***
1<=TTM<=4	Raw Alpha H-L	0.0599	0.21
	CAPM Alpha H-L	-0.0330	-0.12
	Carhart Alpha H-L	-0.0960	-0.41
Maturity-Bucket 2	L	0.8629	2.35*
	2	0.8320	2.94***
	H	0.8400	3.53***
4<TTM<=7	Raw Alpha H-L	-0.0229	0.00
	CAPM Alpha H-L	-0.1590	-0.69
	Carhart Alpha H-L	-0.0730	-0.36
Maturity-Bucket 3	L	0.8700	3.46***
	2	0.8600	2.88***
	H	0.9499	2.47**
7<TTM<=11	Raw Alpha H-L	0.0799	0.02
	CAPM Alpha H-L	-0.1510	-0.62
	Carhart Alpha H-L	-0.1370	-0.60
Maturity-Bucket 4	L	0.8700	3.25***
	2	1.0200	3.47***
	H	0.9678	2.54**
11<TTM	Raw Alpha H-L	0.0978	0.39

	CAPM Alpha H-L	-0.0370	-0.15
	Carhart Alpha H-L	0.0990	0.47

Chapter IV

Affect in a Behavioral Asset Pricing Model

We admire a stock or despise it when we hear its name, whether Google or General Motors, before we think about its price-to-earnings ratio or the growth of its company's sales. Stocks, like houses, cars, watches and most other products exude affect, good or bad, beautiful or ugly, admired or despised. Slovic, Finucane, Peters, and MacGregor (2002) described affect, the specific quality of 'goodness' or 'badness,' as a feeling that occurs rapidly and automatically, often without consciousness. Zajonc (1980), an early proponent of the importance of affect in decision making wrote, "We do not just see house: We see a handsome house, an ugly house, or a pretentious house" (p. 154) and added "We sometimes delude ourselves that we proceed in a rational manner and weigh all the pros and cons of the various alternatives. But this is rarely the case. Quite often 'I decided in favor of X' is no more than "I liked X'. We buy the cars we 'like," choose the jobs and houses we find 'attractive,' and then justify these choices by various reasons." (p. 155) Kahneman (2002) described the affect heuristic in his Nobel Prize Lecture as "probably the most important development in the study of judgment heuristics in the last decades."

Affect plays a role in pricing models of houses, cars and watches but, according to standard financial theory, affect plays no role in pricing of financial assets. Expected returns in the CAPM are determined by risk alone, measured by beta, and, according to Fama and French (1992), market capitalization and book-to-market ratios in their 3-factor asset pricing model of risk. But affect plays a role in behavioral asset pricing models where we know it as ‘sentiment’ or as an ‘expressive’ set of characteristics.

Statman (1999) described a behavioral asset-pricing model that includes utilitarian factors, such as risk, but also expressive or affect characteristics, such as the negative affect of tobacco and other ‘sin’ companies or the positive affect of prestigious hedge funds. He illustrated the model with an analogy to the watch market. A $10,000 Rolex watch and a $50 Timex watch have approximately the same utilitarian qualities; both watches display the same time. But Rolex buyers are willing to pay an extra $9,950 over the price of the Timex because of the affect of a Rolex, consisting of prestige, and perhaps beauty, is more positive than that of a Timex.

Asset pricing models are intertwined with the efficient market hypothesis, but our paper is about asset pricing models, not market efficiency. We find that the returns of stocks admired by respondents of the Fortune surveys were lower than the returns of less admired stocks, but we do not claim to have uncovered a new anomaly. Rather, we hypothesize that affect plays a role in pricing models of financial assets. In particular, we hypothesize that affect underlies the market capitalization and book-to-market factors of the 3-factor models. We find evidence consistent with our hypothesis, and outline a behavioral asset pricing model.

4.1 Affect in pricing models

There is considerable evidence that affect plays a role in pricing. For example, Hsee (1998) presented to subjects pictures of two ice cream cups, depicted in Figure 4.1. The cup of ice cream on the left contains 8 ounces of ice cream but its affect is negative since it seems stingy in its 10-ounce cup. In contrast, the affect of the 7 ounces of ice cream on the right is positive since it is overflowing its 6-ounce cup. Hsee found that subjects who saw only one of the ice cream cups were willing to pay a higher price for the 7 ounces of ice cream with positive affect than for the 8 ounces of ice cream with negative affect. But subjects who saw the two cups side by side were willing to pay a higher price for the cup with 8 ounces of ice cream.

Affect is an emotion and, like all emotions, it is grounded in evolutionary psychology. Cosmides and Tooby (2000) wrote that evolutionary psychology is a theoretical framework that combines principles and results from evolutionary biology, cognitive science, anthropology and neuroscience to describe human behavior. They described emotions as programs whose function is to direct the activities and interactions of sub-programs, including those of perception, attention, goal choice, and physiological reactions. Cosmides and Tooby illustrated with the emotion of fear, as when stalked by predators. "Goals and motivational weightings change; Safety becomes a far higher priority…You are no longer hungry; you cease to think about how to charm a potential mate… adrenalin spikes…" (p.)

Emotions prevent us from being lost in thought when it is time to act. But sometimes emotions subvert good thinking. Reliance on emotions increases with the complexity of information and with stress. Shiv and Fedorikhin (1999) described an experiment where subjects chose between a chocolate cake with intense positive affect but inferior from a cognitive perspective, and a fruit salad with a less positive affect but superior from a cognitive perspective. One group of subjects was assigned a low-stress task, memorizing a two-digit number, while another was assigned a higher-stress task, memorizing a seven-digit number. Next, subjects were asked to walk over to another room. On their way each could choose a chocolate cake or a fruit salad. Shiv and Fedorikhin found that subjects who were under the greater stress of memorizing the seven-digit number were more likely to be guided by affect and choose the chocolate cake over the fruit salad.

Stocks are notoriously complex and their evaluation is stressful. Are shares of Google at $700 per share better investments than shares of General Motors at $20 per share? Investors try to overcome the pull of affect through a systematic examination of relevant information, but affect still exerts its power.

Internet related dotcom names had positive affect in the boom years of the late 1990s and Cooper et al (2001) found that companies that changed their names to dotcom names had positive abnormal returns on the order of 74% in the 10 days surrounding the announcement day, even when nothing about their business has changed. Dotcom names acquired negative affect in the bust years of the early 2000s and Cooper et al (2005) found that companies that changed their dotcom names to conventional names during that time experienced positive abnormal returns once more.

The findings of Cooper et al are examples of 'integral affect.' This is affect that is associated with the characteristics of a particular object, such as a stock. 'Incidental affect' is different from integral affect in that it arises not from an object but from an unrelated event. For example, Welch (1999) induced fear in subjects by showing them two minutes of Kubrick's movie "The Shining." He found that the fear they induced carried over beyond the movie, increasing subjects' risk aversion in choices unrelated to the movie. In the context of stocks, Hirshleifer and Shumway (2003) found that the positive incidental affect of sunny days brought high stock returns, and Edmans et al (2007) found that the negative incidental affect of soccer losses brought low stock returns.

The immediate effect of an increase in affect is an increase in stock prices but higher stock prices set the stage for lower future returns. This long term effect is evident in Hong and Kacperczyk's (2007) study of 'sin' stocks, namely those of tobacco, alcohol and gaming companies. The negative affect of sin companies is reflected in social norms against vice. Hong and Kacperczyk found that stocks of sin companies had abnormal positive returns during the1926 to 2004 time period. We hypothesize that the negative affect of despised companies in the Fortune surveys underlies their higher stock returns, analogous to the higher returns sin company stocks.

4.2 Market efficiency and asset pricing models

Fama (1970) noted that market efficiency per se is not testable. Market efficiency must be tested jointly with an asset pricing model, such as the CAPM or the three-factor model. For example, the excess returns relative to the CAPM of small-cap stocks and stocks with high book-to-market ratios might indicate that the market is not efficient or that the CAPM is a bad model of expected returns. But when it comes to tests of market efficiency the CAPM is quite different from the three-factor model.

The CAPM presents expected returns as a function of objective risk. The objective measure of investment risk is based on the probability distribution of investment outcomes, usually equated with the variance of a portfolio and the beta of a security within a portfolio. In contrast, the three-factor model presents expected returns as functions of beta, a measure of objective risk, but also as functions of market capitalization and book-to-market ratios. But what do market capitalization and book-to-market ratios represent? Fama and French argued that they represent objective risk but much of the evidence is inconsistent with their argument. For example, Lakonishok et al (1994) found that value stocks outperformed growth stocks in three out of four recessions during 1963-1990, inconsistent with the view that value stock are riskier. Similarly, Skinner and Sloan (2002) found that the relatively high returns of value stocks are not due to their higher risk. Rather, they are due to large declines in the prices of growth stocks in response to negative earnings surprises. We present 4-factor analysis of the data here for its insights into assets pricing models, not as a test of market efficiency.

4.3 Fortune admired and despised

Fortune magazine has been publishing the results of an annual survey of company reputations since 1983. The survey published in March 2007 included 587 companies. Fortune asked more than 10,000 senior executives, directors and security analysts who responded to the survey to rate the ten largest companies in their industries on eight attributes of reputation, using a scale of zero (poor) to ten (excellent). We focus on the attribute of Long-Term Investment Value (LTIV) since it reflects perceptions of respondents about company stocks, incorporating both their expected returns and risk.

Consider two portfolios constructed by Fortune scores, each consisting of an equally weighted half of the Fortune stocks. The Admired portfolio contains the stocks of companies with the highest LTIV scores and the Despised portfolio contains the stocks with the lowest scores. If Fortune respondents believe that the stock market is efficient we should expect that they would rate all stock equally on LTIV. This is because in an efficient market there are no stocks with high LTIV and no stocks with low LTIV. If Fortune respondents believe that the stock market is inefficient and they can indeed identify correctly the stocks with higher LTIV, we should expect that stocks of companies with high LTIV would do better than stocks of companies with low LTIV. But this is not what we find. We argue that ratings of LTIV serves as a measure of affect. Fortune respondents rate some stocks high on LTIV and other stocks low because they are influenced by the positive affect of the first group and the negative affect of the other.

We construct the portfolios on September 30, 1982, based on the Fortune survey published subsequently in 1983. This is because Fortune surveys are completed by respondents around September 30th of the year before they are published.

Fortune does not define how long long-term is. We investigate three horizons, 2, 3, and 4 years. For the 2-year horizon we reconstituted each portfolio on September 30th every two years, so the first reconstitution is based on the survey conducted in 1984 and published in 1985. We constructed portfolios similarly for the 3 and 4-year horizons. Fortunately, our overall 24-year period, September 30th 1982 – September 30th 2006 is divisible by all three periods so each time period is included in each analysis.

The mean scores of companies in some industries, such as the 6.43 of the Communication industry, are higher on average than those of other industries, such as the 5.14 of the Coal

Mining industry. We calculate the mean score of companies in each industry in the surveys published in 1983-2007 surveys and define the industry-adjusted score of a company as the difference between its score in a given survey and the mean score of companies in its industry.

The returns of the Despised portfolios exceeded those of the Admired portfolios. For example, the mean annualized return of the Despised portfolio during September 30, 1982 – September 30, 2006 was 19.72% when the portfolio was rebalanced every four years, higher than the 15.12% mean annualized return of the Admired portfolio (see table 4.1)

The advantage of the Despised portfolios over the Admired portfolios remains intact when we assess them by the CAPM. The alphas of the Despised portfolios are consistently higher than those of their respective Admired portfolios. For example, the annualized alpha of the Despised portfolio when portfolios are reconstituted every four years is 4.89% while it is only 1.57% in the Admired portfolio. The alphas of Despised portfolios are positive and statistically significant in all reconstitution intervals. The alphas of the Admired portfolios are always positive but statistically significant only in the 3-year reconstitution interval.

4.4 Characteristics of despised and admired portfolios

A 4-factor analysis, presented in Table 4.2, shows that companies in the Despised portfolios have higher objective risk than companies in the Admired portfolios. Betas in the Despised portfolios are consistently higher than betas in the respective Admired portfolios. The 4-factor analysis also shows that the characteristics of small, value and low short-term momentum are associated with the Despised portfolios. The tilts of the Despised portfolios toward small and value are consistently greater than those of the respective Admired portfolios and the momentum of the Despised portfolios is consistently lower than that of the Admired portfolios. Further analysis presented in Table 4.3 shows that companies in the Despised portfolios also had higher earnings-to-price ratios, higher cash-flows-to-price ratios, lower past sales and earnings growth and lower returns on assets.

4.5 Affect in a behavioral asset pricing model

The behavioral asset pricing model we outline is one where expected returns are high when objective risk is high and also when subjective risk is high. High subjective risk comes with negative affect and low subjective risk comes with positive affect.

Subjective risk is different from objective risk. For example, Ganzach (2000) presented a list of 30 international stock markets to two groups of subjects. One group was asked to judge the expected returns of the market portfolios of each stock market, while the other group was asked to judge the risk of these market portfolios. A CAPM-like asset pricing model based entirely on objective risk would lead us to expect a positive correlation between assessments of risk and assessments of expected returns but Ganzach found a negative correlation; markets with high expected returns were perceived to have low risk.

The negative relationship between subjective risk and expected returns in Ganzach's study is one example of a general negative relationship between subjective risk and perceived benefits. Slovic et al (2002) attribute that negative relationship to the halo of affect. When affect is positive benefits are judged high and risk is judged low. And when affect is negative benefits are judged low and risk high. We find similar results in our experiments.

In the first experiment, conducted in May 2007, we asked investors, high net-worth clients of an investment company, to complete a questionnaire listing only the names of 210 companies from the Fortune 2007 survey, their industries, and a 10-point scale ranging from "bad" to "good". The questionnaire said: "Look at the name of the company and its industry and quickly rate the feeling associated with it on a scale ranging from bad to good. Don't spend time thinking about the rating. Just go with your quick, intuitive feeling." The affect score of a company is the mean score assigned to it by the surveyed investors. We found a positive and statistically significant relationship between affect scores and Fortune scores (see Figure 4.2).

In the second experiment, conducted in July 2007 we presented to another group of investors the names and industries of the same 210 companies from the Fortune 2007 survey. One group of investors was asked to rate the future return of each stock on a 10-point scale ranging from low to high. Another group of investors was asked to rate the risk of each stock on the same scale. The risk and return scores of companies are the mean scores assigned to them by the surveyed investors.

If investors' assessment of risk reflects objective risk alone we should find a positive correlation between the risk scores and the return scores they assigned to companies. However, as seen in Figure 4.3, we find a negative correlation between the two; high return scores correspond to low risk score. This negative correlation indicates that investors assessments of risk reflect subjective risk associated with affect. Affect creates a halo over stocks. Stocks with positive affect are assessed high in future returns and low in risk, and stocks with negative affect are assessed low in future returns and high in risk.

We also find a link between return scores, risk scores, and Fortune scores. In a regression of Fortune scores on return scores we find that high Fortune ratings are associated with high return scores. The coefficient of the return scores is positive and statistically significant. Similarly, in a regression of Fortune scores on risk scores we find that high Fortune ratings are associated with low risk scores. The coefficient of the risk scores is negative and statistically significant (see Figures 4.4 and 4.5).

Objective risk measured by beta and subjective risk measured by affect are two factors in the behavioral asset pricing model. But they are not alone. Momentum is an especially interesting factor since its rationale is distinct from the rationale of affect.

Objective risk measured by beta and subjective risk measured by affect are two factors in the behavioral asset pricing model. But they are not alone. Short-term momentum is an especially interesting factor since its rationale is distinct from the rationale of affect.

Short-term (12-month) momentum is positively correlated with affect, yet it is generally associated with high returns (Jagadeesh and Titman (1993)). In contrast, market capitalization which is also positively correlated with affect is generally associated with low returns. This suggests that the association between short-term momentum and returns is not due to the role of short-term momentum as a proxy for affect. Indeed, the association between short-term momentum and returns has been attributed by Grinblatt and Han (2005) to the "disposition effect," described by Shefrin and Statman (1985) and by Sias (2007) to trading by institutional investors.

4.6 Investor preferences and stock returns

The road from the perception that admired companies offer both high expected returns and low risk to the low realized returns of such stocks is not straight, as explained by Shefrin and Statman (1995) and more recently by Pontiff (2006). Suppose that typical investors prefer admired companies they perceive as having both high expected returns and low risk. But surely some investors are 'contrarians,' aware of the preferences of typical investors and seek capitalize on them by favoring stocks of despised companies. Would arbitrage by contrarians not nullify any effect of typical investors on stock returns? Subjective risk stemming from affect plays no role in the asset pricing model if the effects of typical investors on stock returns are nullified by arbitrage. However, subjective risk plays a role in the asset pricing model if arbitrage is incomplete.

As we consider arbitrage and the likelihood that it would nullify the effects of the preferences of typical investors on stock returns we should note that no perfect (risk-free) arbitrage is possible here. As some hedge funds and other unlucky investors found out, price gaps that are likely to close over a long period might widen further over a shorter period. To see the implications of imperfect arbitrage, imagine contrarians who know that stocks of despised companies have high expected returns relative to their objective risk. It is optimal for contrarians to increase their holdings of stocks of despised companies, but as the amount devoted to such stocks increases, the portfolios of contrarians become less diversified and they take on more idiosyncratic risk. The increase in portfolio risk leads contrarians to limit the amount allocated to despised stocks, and with it, limit their effect on stock returns.

4.7 Conclusion

All asset pricing models, whether of securities, cars or watches, are versions of the basic demand and supply model where prices are determined by the intersection of demand and supply. The demand and supply functions reflect the preferences of consumers and producers.

The demand and supply structure is evident in the CAPM. In that model investors on both the demand and supply sides prefer mean-variance-efficient portfolios and the aggregation of their preferences yields an asset pricing model where expected returns of securities vary by beta. The demand and supply structure is not nearly as evident in the Fama and French 3-factor asset pricing model. Market capitalization and book-to-market ratios were associated with anomalies relative to the CAPM long before their debut in the 3-factor model, but the argument that market capitalization and book-to-market ratios proxy for risk is not fully supported by the evidence.

The purpose of this paper is to help link asset pricing models to the preferences of investors. We outline a behavioral asset pricing model where expected returns are high when objective risk is high and also when subjective risk is high. High subjective risk comes with negative affect and low subjective risk comes with positive affect. Affect is the specific quality of ‘goodness’ or ‘badness.’ It is a feeling that occurs rapidly and automatically, often without consciousness. Investors prefer stocks with positive affect and their preference boosts the prices of stocks with positive affect and depresses their returns.

We study the preferences of investors as reflected in surveys conducted by Fortune magazine during 1983- 2006 and additional surveys we conducted in 2007. We find that the returns of admired stocks, those highly rated by the Fortune respondents, were lower than the returns of despised stocks, those rated low. This is consistent with the hypothesis that stocks with negative affect have high subjective risk and their extra returns compensate for that risk. We also find that market capitalization and book-to-market ratios are correlated with affect and argue that they proxy for it.

We find additional evidence consistent with the hypothesis in our own surveys. Respondents in our surveys rate companies as if they believe that stocks with high expected returns also have low risk and perceive stocks of companies admired by Fortune respondents as having both high expected returns and low risk.

We emphasize that the behavioral asset pricing model we outline is not superior to the 3 or 4-factor models. Indeed, the 3 and 4-facor models are behavioral models under their standard-finance skins. The affect factor in the behavioral asset pricing model elucidated the rationale underlying the market cap and book-to-market factors of the 3-factor model. The number of factors in a full model is likely to grow to include factors such as liquidity that are not included in our behavioral model or in the 3 and 4-factor models. Moreover, affect has several distinct sources and these sources might play distinct roles in a behavioral asset pricing model. Social responsibility is one source of positive affect, and tobacco companies lack it. Prestige is another source of positive affect, and hedge funds posses it.

Table 4.1: CAPM-based performance of Admired and Despised portfolios

September of every second, third and fourth year from 1982 to 2006, we form two portfolios based on the 'overall reputation' score of each company in the *Fortune* survey. The 'Admired' portfolio contains all the companies ranked highest and the 'Despised' portfolio contains all the companies ranked lowest based on their scores each year. In forming portfolios, we adjust the overall company scores for industry differences. The adjusted score is computed by subtracting the average industry score from each individual company's score for the industry to which the company belongs. The first two digits of the SIC code is used for industry classification in computing the average industry scores We then calculate equally weighted returns for the two portfolios from March of year t to September of year t+2, t+3 and t+4. Companies delisted during the holding period are assigned their delisting return and removed from the portfolio next month. In this table we report CAPM alphas for the two portfolios for the t+2, t+3 and t+4 holding periods . *t-statistics* are reported in below coefficient estimates. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Portfolios reconstituted every 2 years.

	Despised	**Admired**	**Difference**
Mean annualized return	18.99%	15.65%	3.34%
	CAPM-Based Performance[2]		
Annualized Alpha	4.37%	1.94%	2.43%
t-stat	2.43**	1.67*	
Market	1.04	0.98	0.06
t-stat	30.84***	44.82***	
Adj R^2	0.76	0.87	

Portfolios reconstituted every 3 years.

	Despised	**Admired**	**Difference**
Mean annualized return	17.83%	16.02%	1.81%
	CAPM-Based Performance[2]		
Annualized Alpha	3.81%	2.29%	1.52%
t-stat	2.17**	1.95*	
Market	1.03	1.00	0.04
t-stat	31.24***	44.58***	
Adj R^2	0.77	0.87	

Portfolios reconstituted every 4 years.

	Despised	**Admired**	**Difference**
Mean annualized return	19.72%	15.12%	4.60%
	CAPM-Based Performance[2]		
Annualized Alpha	4.89%	1.57%	3.32%

t-stat	2.82***	1.31	
Market	1.03	0.98	0.05
t-stat	31.66***	42.96***	
Adj R^2	0.77	0.86	

Table 4.2: 4-factor-based performance of Admired and Despised portfolios

September of every second, third and fourth year from 1982 to 2006, we form two portfolios based on the 'overall reputation' score of each company in the *Fortune* survey. The 'Admired' portfolio contains all the companies ranked highest and the 'Despised' portfolio contains all the companies ranked lowest based on their scores each year. In forming portfolios, we adjust the overall company scores for industry differences. The adjusted score is computed by subtracting the average industry score from each individual company's score for the industry to which the company belongs. The first two digits of the SIC code is used for industry classification in computing the average industry scores We then calculate equally weighted returns for the two portfolios from March of year t to September of year t+2, t+3 and t+4. Companies delisted during the holding period are assigned their delisting return and removed from the portfolio next month. In this table we report 4-factor alphas for the two portfolios for the t+2, t+3 and t+4 holding periods . *t-statistics* are reported in below coefficient estimates. Statistical significance at the 10%, 5% and 1% levels is denoted by *, **, and ***, respectively.

Portfolios reconstituted every 2 years.

	Despised Portfolio	Admired Portfolio	Difference
	4-Factor Based Performance[2]		
Annualized Alpha	1.90%	0.35%	1.55%
t-stat	1.55	0.36	
Market	1.18	1.09	0.09
t-stat	45.75***	53.61***	
Small-minus-Big	0.36	-0.05	0.41
t-stat	11.25***	-1.99***	
Value-minus-Growth	0.59	0.29	0.29
t-stat	15.26***	9.66***	
Momentum	-0.24	-0.09	-0.15
t-stat	-10.60***	-4.95***	
Adj R^2	0.90	0.92	

Portfolios reconstituted every 3 years.

	Despised Portfolio	Admired Portfolio	Difference
	4-Factor Based Performance[2]		
Annualized Alpha	1.29%	0.81%	0.48%
t-stat	1.04	0.83	
Market	1.17	1.10	0.06
t-stat	44.60***	54.08***	
Small-minus-Big	0.35	-0.04	0.39
t-stat	10.81***	-1.46	
Value-minus-Growth	0.57	0.30	0.26
t-stat	14.54***	9.95***	
Momentum	-0.22	-0.11	-0.11
t-stat	-9.53***	-5.94***	
Adj R^2	0.89	0.92	

Portfolios reconstituted every 4 years.

	Despised Portfolio	Admired Portfolio	Difference
	4-Factor Based Performance[2]		
Annualized Alpha	2.07%	-0.02%	2.09%

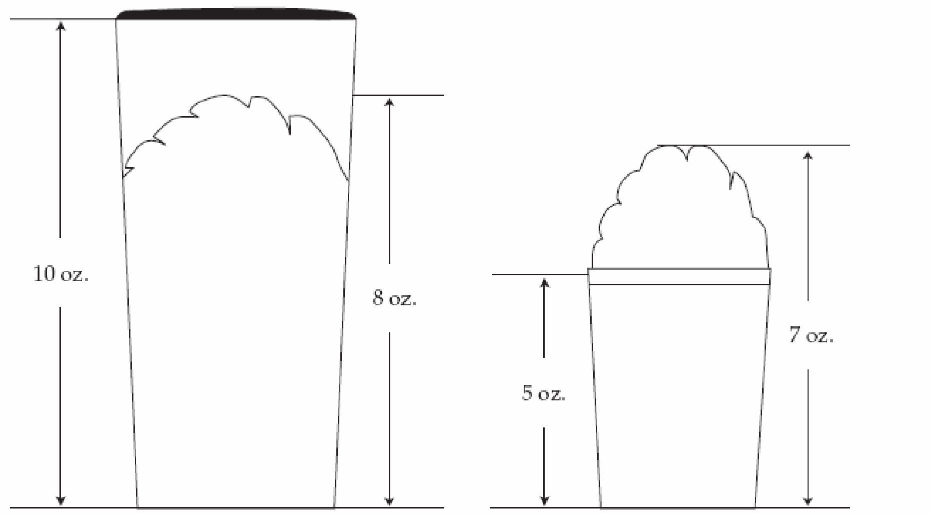
t-stat	1.64	-0.03	
Market	1.17	1.09	0.08
t-stat	44.18***	52.01***	
Small-minus-Big	0.32	-0.02	0.34
t-stat	9.70***	-0.96	
Value-minus-Growth	0.57	0.32	0.25
t-stat	14.42***	10.11***	
Momentum	-0.19	-0.11	-0.09
t-stat	-8.25***	-5.72***	
Adj R^2	0.89	0.92	

Table 4.3: Characteristics of stocks in admired and despised portfolios

September of each year, we form two portfolios based on the 'overall reputation' score of each company in the *Fortune* survey. The 'Admired' portfolio contains all the companies ranked highest and the 'Despised' portfolio contains all the companies ranked lowest based on their scores each year. In forming portfolios, we adjust the overall company scores for industry differences. The adjusted score is computed by subtracting the average industry score from each individual company's score for the industry to which the company belongs. The first two digits of the SIC code is used for industry classification in computing the average industry scores. In this table we report average characteristics of stocks in each portfolio. 1Market capitalization is at the end of September of the portfolio formation year. Book equity (defined as in Davis, Fama, French 2000) at the end of the fiscal year prior to portfolio formation and price at the end of September of the portfolio formation year. Earnings are in the fiscal year prior to portfolio formation and price at the end of September of the portfolio formation year. Cash flow (Earnings + Depreciation) in the fiscal year prior to portfolio formation and price at the end of September of the portfolio formation year. These ratios are set to zero if they are negative. Sales growth is log change in sales in the two fiscal years prior to the end of September of the portfolio formation year. Earnings growth is log change in earnings in the two fiscal years prior to the end of September of the portfolio formation year. Return on Assets (ROA) is calculated as the ratio of operating income before depreciation to total assets at the end of the fiscal year

	Mean Values as of September 30 of each year, 1982 - 2005.	
	Stocks in the Admired Portfolio	**Stocks in the Despised Portfolio**
Returns in the previous year	21.57%	11.06%
Returns in the previous 3 years	81.24%	38.47%
Returns in the previous 5 years	169.44%	79.50%
Market Capitalization ($ millions)[1]	19,327	5,853
Book-to-Market ratio	0.491	0.751
Earnings-to-Price ratio	0.066	0.079
Cash-Flow-to-Price ratio	0.103	0.136
Sales Growth	0.101	0.035
Earnings Growth	0.127	0.052
Return on Assets	0.158	0.125
Beta	0.980	1.040

Figure 4.1: Hsee (1998) experiment



Source: Hsee (1998).

Figure 4.2: The relationship between affect scores and Fortune scores

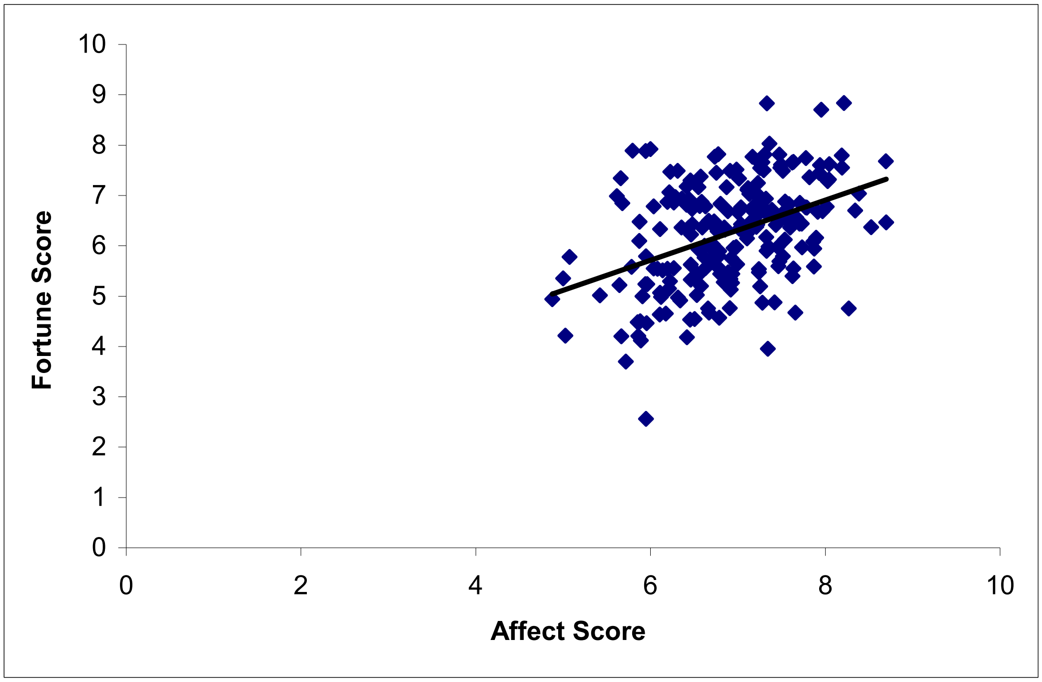


Figure 4.3: The relationship between expected return scores and risk scores

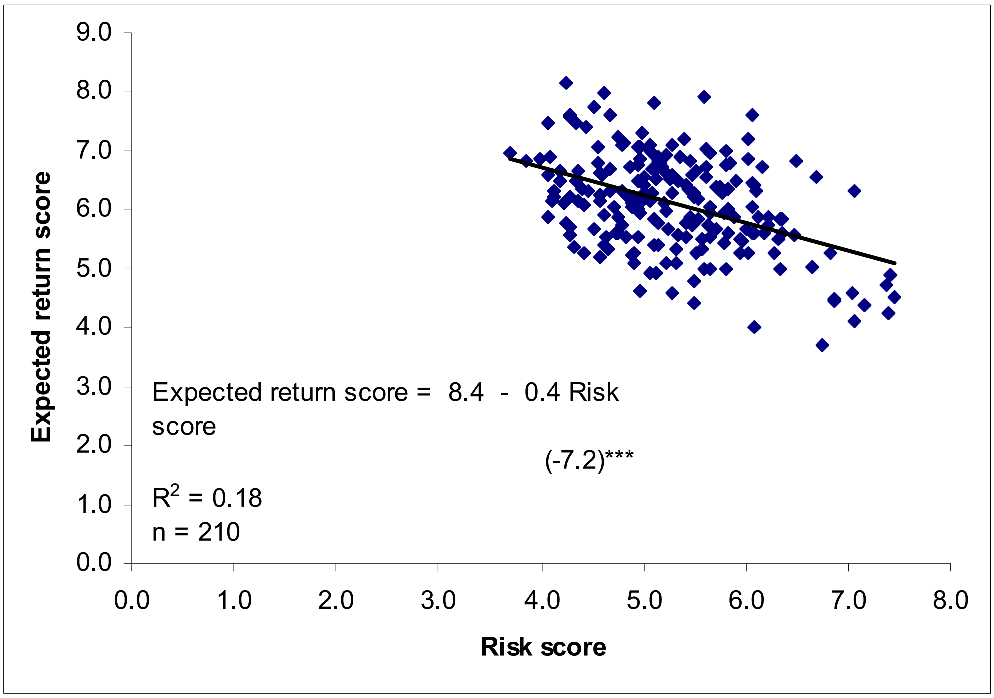


Figure 4.4: The relationship between expected return scores and Fortune scores

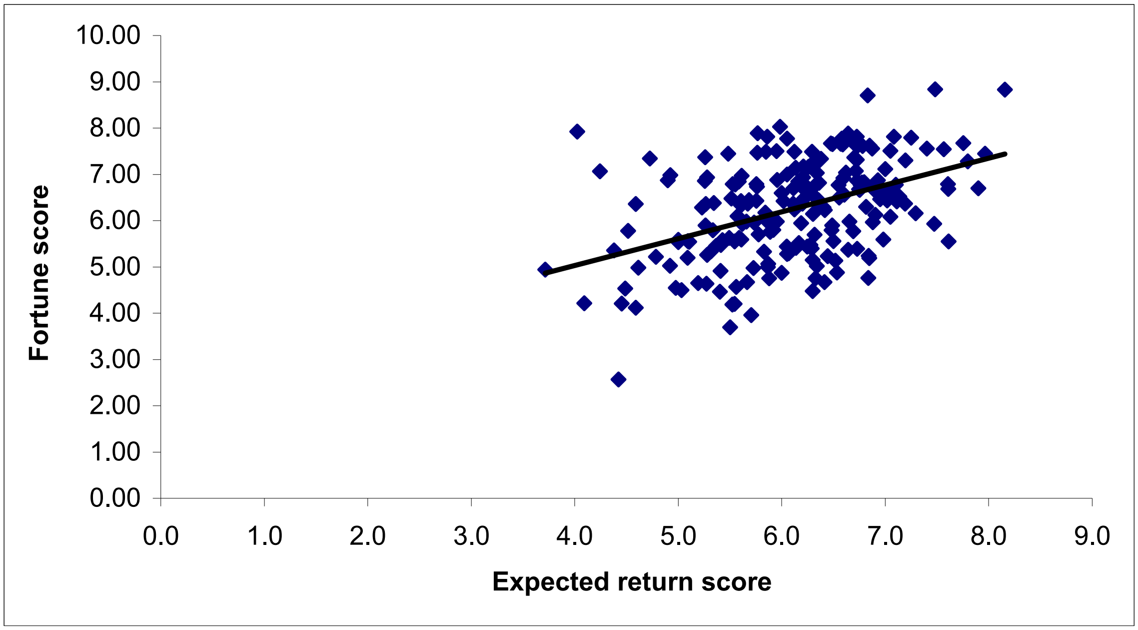
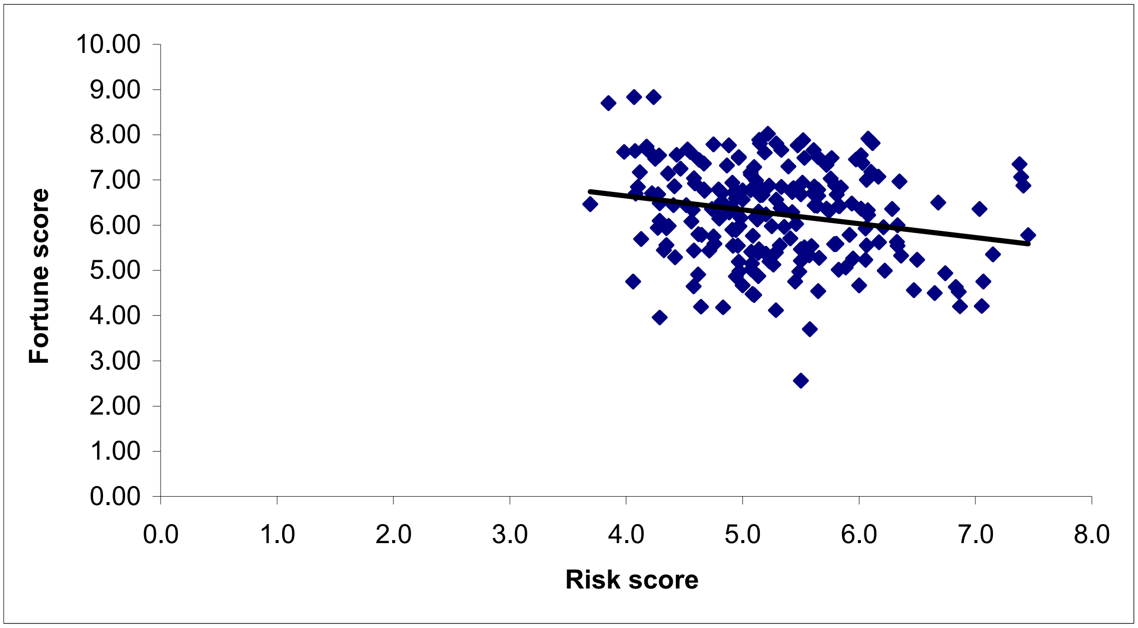


Figure 4.5: The relationship between risk scores and Fortune scores



Appendix

Here we explain the details of the variables used to construct distress measures. Quarterly COMPUSTAT data is used to compute all accounting variables. Our first measure is Altman z-score, which is defined as the following:

$$\text{z-score} = 1.2\ WCTA + 1.4\ RETA + 3.3\ EBITTA + 0.6\ METL + 1.0\ STA \quad \text{(A1)}$$

WCTA is the working capital (data40 – data49) divided by total assets. We follow CHS 2008 to adjust total assets calculated as total liabilities (data54) + market equity + 0.1*(market equity – book equity). Book equity is as defined in Davis, Fama, and French (2000). *RETA* is the ratio of retained earnings (data58) to total assets. *EBITTA* is the ratio of earnings before interest and taxes (data21 - data5 + data31) to total assets, *METL* is the ratio of market equity to total liabilities, and *STA* is the ratio of sales (data12) to total assets. Our second measure is Ohlson's o-score, defined as:

$$\begin{aligned} o\text{-score} &= -1.32 - 0.407\log(SIZE) + 6.03\ TLTA - 1.43\ WCTA \\ &+0.076\ CLCA - 1.72\ OENEG - 2.37\ NITA - 1.83\ FUTL \qquad \text{(A2)} \\ &+0.285\ INTWO - 0.521\ CHIN \end{aligned}$$

where *SIZE* is total assets divided by the consumer price index, *CLCA* is the ratio of current liabilities (data49) to current assets (data40), *OENEG* is a dummy variable equal to one if total liabilities exceeds total assets and zero otherwise, *NITA* is the ratio of net income (data69) to total assets, *FUTL* is the ratio of funds from operations (data23) to total liabilities, *INTWO* is a dummy variable equal to one if net income was negative for the past two years and zero otherwise, and *CHIN* is change in net income over the last quarter: $(NI_t NI_{t-1})/(|NI_t| + |NI_{t-1}|)$.

The third measure we use is the CHS-score:

$$\begin{aligned} CHS\text{-}score_t \;=\; & -9.164 - 20.264\ NIMTAAVG_t + 1.416\ TLMTA_t \\ & -7.129\ EXRETAVG_t + 1.411\ SIGMA_t - 0.045\ RSIZE_t \\ & -2.132\ CASHMTA_t + 0.075\ MB_t - 0.058\ PRICE_t \end{aligned} \qquad \text{(A3)}$$

where *NIMTAAVG* is a geometrically declining average of past values of the ratio of net income (data69) to total assets:

$$NIMTAAVG_{t-1,t-12} = \frac{1-\phi^2}{1-\phi^{12}}\left(NIMTA_{t-1,t-3} + ... + NIMTA_{t-10,t-12}\right) \qquad \text{(A4)}$$

EXRETAVG is a geometrically declining average of monthly log excess stock returns relative to the S&P 500 index:

$$EXRETAVG_{t-1,t-12} = \frac{1-\phi}{1-\phi^{12}}\left(EXRET_{t-1} + ... + \phi^{11}EXRET_{t-12}\right) \qquad \text{(A5)}$$

The weighting coefficient is set to $\phi = 2^{-1/3}$, such that the weight is halved each quarter. *TLMTA* is the ratio of total liabilities (data69) to total assets. *SIGMA* is the standard deviation of daily stock returns over the previous three months. *SIGMA* is coded as missing if there are fewer than 5 observations. *RSIZE* is the log ratio of market capitalization to the market value of the S&P 500 index. *CASHMTA* is the ratio of the value of cash and short term investments (data36) to the value of total assets. *MB* is the market-to-book ratio. Book equity is as defined in Davis, Fama, and French (2000). *PRICE* is the log price per share truncated from above at \$15. All variables are winsorized using a 5/95 percentile interval in order to eliminate outliers.

We follow CHS (2008) and Hillegeist et al. (2004) to calculate our fourth distress measure, Merton's distance-to-default. The market equity value of a company is modeled as a call option on the company's assets:

$$\begin{aligned} V_E &= V_A e^{-\partial T} N(d_1) - Xe^{-rT} N(d_2) + (1 - e^{-\partial T})V_A \\ d_1 &= \frac{\log(V_A / X) + (r - \partial - (\sigma_A^2 / 2))T}{\sigma_A \sqrt{T}} \\ d_2 &= d_1 - \sigma_A \sqrt{T} \end{aligned} \qquad \text{(A6)}$$

Above V_E is the market value of a firm. V_A is the value of firm's assets. X is the face value of debt maturing at time T. r is the risk-free rate and ∂ is the dividend rate expressed in terms of V_A. σ_A is the volatility of the value of assets, which is related to equity volatility through the following equation:

$$\sigma_E = \left(V_A e^{-\partial T} N(d_1) \sigma_A \right) / V_E \qquad \text{(A7)}$$

We simultaneously solve the above two equations to find the values of V_A and σ_A. We use the market value of equity for V_E and short-term plus one half long-term book debt to proxy for the face value of debt X (data45+1/2*data51). σ_E is the standard deviation of daily equity returns over the past 3 months. T equals one year, and r is the one-year treasury bill rate. The dividend rate, *d*, is the sum of the prior year's common and preferred dividends (data19 + data21) divided by the market value of assets. We use the Newton method to simultaneously solve the two equations above. For starting values for the unknown variables we use, $V_A = V_E + X$, and $\sigma_A = \sigma_E V_E (V_E + X)$. Once we determine asset values, V_A, we then compute asset returns as in Hillegeist et al. (2004):

$$\mu_t = \max\left[\frac{V_{A,t} + Dividends - V_{A,t-1}}{V_{A,t-1}}, r \right] \qquad \text{(A8)}$$

As expected returns cannot be negative, if asset returns are below zero they are set to the risk-free rate.[44] Merton's distance-to-default is finally computed as:

$$MertonDD = -\frac{\log\left(V_A / X \right) + \left(\mu - \partial - \left(\sigma_A^2 / 2 \right) \right) T}{\sigma_A \sqrt{T}} \qquad \text{(A9)}$$

[44] We obtain similar results if we use a 6% equity premium instead of asset returns as in CHS (2008).

Bibliography

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

Almeida, H. and T. Philippon, 2007, "The risk-adjusted cost of financial distress," Journal of Finance, 62, 2557-2586.

Altman, Edward I., 1968, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," Journal of Finance, 23, 589—609.

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.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaozan Zhang, 2006, "The cross-section of volatility and expected returns," Journal of Finance, 61, 259—299.

Ang A., Hodrick R., Xing Y., Zhang X, 2009, "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence," Journal of Financial Economics, 91, 1–23.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2009, "Credit Ratings and The Cross-Section of Stock Returns," Journal of Financial Markets, 123, 469-499.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2006, "Momentum and Credit Rating," Journal of Finance, 62, 2503-2520.

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

Bakshi, G., Madan, D., Zhang, F., 2006, "Investigating the role of systematic and firm-specific factors in default risk: Lessons from empirically evaluating credit risk models," Journal of Business, 79, 1955–1988.

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, “Systemic 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.

Barberis, Nicholas, and Ming Huang, 2001, “Mental accounting, loss aversion, and individual stock returns,” Journal of Finance, 56, 1247–1292.

Berndt, A., Douglas, R., Duffie, D., Ferguson, M. and D. Schranz, 2005, “Measuring Default Risk Premia from Default Swap Rates and EDFs,” Working Paper, Stanford University.

Bharath, Sreedhar and Tyler Shumway, 2008, “Forecasting default with the KMV Merton model,” Review of Financial Studies, 213, 1339-1369.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, “Alternative Factor Specifications, Security Characteristics, and the Cross-Secion of Expected Stock Returns,” Journal of Financial Economics, 49, 345–373.

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, John Y., Jens Hilscher, and Jan Szilagyi, 2008, “In Search of Distress Risk,” Journal of Finance, 63, 2899-2939.

Campbell, John Y., and Glen B. Taksler, 2003, “Equity Volatility and Corporate Bond Yields,” Journal of Finance, 58, 2321–2350.

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

Campello, Murillo, Lu Chen and Lu Zhang 2008, “expected returns, yield spreads, and asset pricing tests,” Review of Financial Studies, 21, 1297 1338.

Carhart, Mark, 1997, “On Persistence in Mutual Fund Performance,” Journal of Finance, 521, 57–82.

Chava, Sudheer and Robert A. Jarrow, 2004, “Bankruptcy prediction with industry effects,” Review of Finance, 8, 537—569.

Chen, L., Lesmond, D. A. & Wei, J., 2005, “Corporate yield spreads and bond liquidity,” Journal of Finance, 62, 119–149.

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.

Cochrane, John H., 1991, “Production-based asset pricing and the link between stock returns and economic fluctuations”, Journal of Finance, 46, 209–237.

Cooper, Michael J., Raghavendra Rau and Orlin Dimitrov, 2001, "A Rose.com by any other Name," The Journal of Finance, 56, 2371-2388.

Cooper, Michael J., Raghavendra Rau, Ajay Patel, Igor Osobov, and Ajay Khorana, 2005, “Managerial actions in response to a market downturn, Corporate name changes during the dot.com decline,” The Journal of Corporate Finance, 11, 319-335.

Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin, 2001, “The determinants of credit spread changes,” Journal of Finance, 56, 2177–2207.

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.

Cosmides, Leda and John Tooby, 2000, “Eutionary Psychology and the Emotions,” taken from Handbook of Emotions, 2nd Edition, Lewis and Jones eds, 91-115.

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.

Da, Zhi and Pengjie Gao, 2004, "Default risk and equity return: macro effect or micro noise?" Working paper, Northwestern University

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

Das, S.R., Duffie, D., Kapadia, N., Saita, L., 2006, "Common failings: how corporate defaults are correlated," Working paper, Stanford University, forthcoming, Journal of Finance.

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

Davis, J.L., E.F. Fama and K.R. French, 2000, "Characteristics, Covariances, and Average Returns, 1929 to 1997," Journal of Finance, 551, 389-406.

Dichev, Ilia D., 1998, "Is the Risk of Bankruptcy a Systematic Risk?" Journal of Finance, 533, 1131–1147.

Driessen, J., 2005, "Is default event risk priced in corporate bonds?" Review of Financial Studies, 181, 165-195.

Driessen, J. and Frank de Jong, 2007, "Liquidity Risk Premia in Corporate Bond Markets," Management Science, 539, 1439-1451.

Duffee, Gregory, 1999, "Estimating the price of default risk," Review of Financial Studies, 12, 197– 226.

Duffie, Darrell, and Kenneth J. Singleton, 1995, "Modeling term structures of defaultable bonds," Working paper, Stanford Graduate School of Business.

Duffie, Darrell, and Kenneth J. Singleton, 1997, "An econometric model of the term structure of interest-rate swap yields," Journal of Finance, 52, 1287–1321.

Duffie, D., Saita, L., Wang, K., 2007, "Multi-period corporate default prediction with stochastic covariates," Journal of Financial Economics, 833, 635-665.

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.

Edmans, Alex, D. Garcia and O. Norli, 2007, "Sports Sentiment and Stock Returns," Journal of Finance, 62,1967-1998.

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann, 2001, "Explaining the Rate Spread on Corporate Bonds," Journal of Finance, 561, 247–277.

Eom, Young Ho, Jean Helwege, and Jing-Zhi Huang, 2004, "Structural Models of Corporate Bond Pricing: An Empirical Analysis," Review of Financial Studies, 17, 499-535.

Falkenstein, Eric G., 1996, "Preferences for Stock Characteristics as Revealed by Mutual Fund Portfolio Holdings," Journal of Finance, 511, 111-135.

Fama, Eugene, 1970, "Efficient Capital Markets, A Review of Theory and Empirical Work," Journal of Finance, 25, 383-417.

Fama, Eugene F., and Kenneth R. French, 1992, "The Cross-Section of Expected Stock Returns," Journal of Finance, 472, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," Journal of Financial Economics, 33, 3–56.

Fama, Eugene F., and James D. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," Journal of Political Economy, 81, 607–636.

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.

Ferguson, Michael F. and Richard L. Shockley, 2003, "Equilibrium anomalies," Journal of Finance, 58, 2549—2580.

Ganzach, Yoav, 2000, "Judging risk and return of financial assets," Organizational Behavior and Human Decision Processes, 83, 353-370.

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

Garlappi, Lorenzo, Tao Shu, and Hong Yan, 2008, "Default Risk, Shareholder Advantage, and Stock Returns," Review of Financial Studies, 216, 2743-2778.

George, Thomas J. and Hwang, Chuan-Yang, 2009, “A Resolution of the Distress Risk and Leverage Puzzles in the Cross Section of Equity Returns,” forthcoming, Journal of Financial Economics.

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.

Griffin, John M. and Michael L. Lemmon, 2002, “Book-to-market equity, distress risk, and stock returns,” Journal of Finance, 57, 2317—2336.

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.

Hillegeist, Stephen A., Elizabeth Keating, Donald P. Cram and Kyle G. Lunstedt, 2004, “Assessing the probability of bankruptcy,” Review of Accounting Studies, 9, 5—34.

Hirshleifer, David and Tyler Shumway, 2003, "Good day sunshine, Stock returns and the weather," Journal of Finance, 583, 1009-1032.

Hong, Harrison and Marcin Kacperczyk, 2007, “The price of sin: The effects of social norms on the market,” Princeton University, Working paper.

Hsee, C. K., 1998, “Less is better, When low-value options are judged more highly than high-value options,” Journal of Behavioral Decision Making, 11, 107-21.

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

Huang, Jing-Zhi, and Ming Huang, 2003, “How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk?” Working paper, Pennsylvania State University.

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

Hull, J., Predescu, M., White, A., 2003, “The relationship between credit default swap spreads, bond yields, and credit rating announcements,” Journal of Banking and Finance, 2811, 2789-2811.

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.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” Journal of Finance, 481, 35–91.

Jones, Charles M., and Matthew Rhodes-Kropf, 2003, “The price of diversifiable risk in venture capital and private equity,” Working paper, Columbia University.

Kahneman, Daniel, 2002, “Maps of bounded rationality, A perspective on intuitive judgment and choice,” Nobel Prize lecture, December 8.

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.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, “Contrarian Investment, Extrapolation, and Risk,” The Journal of Finance, 42, 1541-1578.

Li, E.X., Livdan, D., Zhang, L., 2007, “Anomalies,” Review of Financial Studies, 2211, 4301-4334.

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.

Lintner, John, 1965, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," Review of Economics and Statistics, 47, 13– 37.

Liu, Laura Xiaolei, Toni M. Whited and Lu Zhang, 2009, "Investment-based expected stock returns," Working Paper, University of Michigan, forthcoming, Journal of Political Economy.

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.

Longstaff, Francis A., and Eduardo S. Schwartz, 1995, "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt," Journal of Finance, 503, 789–821.

Longstaff, F., Mithal, S., Neis, E., 2005, "Corporate yield spreads: default risk or liquidity? New evidence from the credit-default swap market," Journal of Finance, 605, 2213-2253

Malkiel, Burton G., and Yexiao Xu, 2002, "Idiosyncratic risk and security returns," Working paper, University of Texas at Dallas.

Merton, Robert C., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," Journal of Finance, 29, 449–470.

Merton, Robert C., 1987, "Presidential address: A simple model of capital market equilibrium within complete information," Journal of Finance, 42, 483–510.

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.

Newey, Whitney, and Kenneth West, 1987, "A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix," Econometrica, 55, 703–708

Odean, Terrance, 1999, "Do investors trade too much?," American Economic Review, 89, 1279–1298.

Ohlson, James A., 1980, "Financial ratios and the probabilistic prediction of bankruptcy," Journal of Accounting Research, 18, 109—131.

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

Pastor, Lubos, and Pietro Veronesi, 2003, “Stock Valuation and Learning about Profitability,” Journal of Finance 585), 1749–1790

Penman, S., S. Richardson, and I. Tuna, 2007, “The book-to-price effect in stock returns: accounting for leverage,” Journal of Accounting Research, 45, 427-467.

Pontiff, Jeffrey, 2005, “Costly arbitrage and the myth of idiosyncratic risk,” Journal of Accounting and Economics, October, 35-52.

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

Saita, L, 2006, “The Puzzling Price of Corporate default Risk,” Working Paper, Stanford University.

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

Sharpe, William F., 1964, “Capital Asset Prices: A Theory of Market Equilibrium,” Journal of Finance, 19, 425–442.

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.

Shefrin, H., and Meir Statman, 1995, “Making sense of beta, size and book to market,” Journal of Portfolio Management, 19, 75-98.

Shiv, Baba and Alexander Fedorikhin, 1999, “Heart and mind in conflict, The interplay of affect and cognition in consumer decision making,” The Journal of Consumer Research, 26, 278-292.

Shumway, Tyler, 2001, “Forecasting bankruptcy more accurately: a simple hazard model,” Journal of Business, 74, 101—124.

Slovic, Paul, Melissa Finucane, Ellen Peters and Donald G. MacGregor, 2002, “The affect heuristic,” Heuristics and Biases, Gilovich, Griffin and Kahneman eds, New York, Cambridge University Press.

Skinner, Douglas J. and Richard G. Sloan, 2002, “Earnings Surprises, Growth Expectations, and Stock Returns or Don't Let an Earnings Torpedo Sink Your Portfolio,” Review of Accounting Studies, 7, 289-211.

Statman, Meir, 1999, “Behavioral Finance, Past battle and future engagements,” Financial Analysts Journal, November/December, 18-27.

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.

Vassalou, Maria and Yuhang Xing, 2004, "Default risk in equity returns," Journal of Finance, 59, 831—868.

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.

Welch, Ned, 1999, "The heat of the moment," Doctoral Dissertation, Department of Social and Decision Sciences, Carnegie Mellon University.

Zajonc, R.B, 1980, "Feeling and thinking, Preferences need no inferences," American Psychologist, 35, 151-175.

Zhang, Lu, 2007, "Discussion: In Search of Distress Risk," Conference on Credit Risk and Credit Derivatives Federal Reserve Board.

Zhu, H., 2004, "An empirical comparison of credit spreads between the bond market and the credit default swap market," BIS Working Paper No. 160.

Zmijewski, Mark E., 1984, "Methodological issues related to the estimation of financial distress prediction models," Journal of Accounting Research, 22, 59—82.