# INDIVIDUAL INVESTING IN THE FAMILIAR AND THE UNFAMILIAR

By

ABHISHEK VARMA

A dissertation/thesis submitted in partial fulfillment of
the requirement for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
College of Business

MAY 2009

To the faculty of Washington State University:

The members of the Committee appointed to examine the dissertation/thesis of ABHISHEK VARMA find it satisfactory and recommend that it be accepted.

____________________________________

John R. Nofsinger, Ph. D., Chair

____________________________________

Richard W. Sias, Ph. D.

____________________________________

Swaminathan Kalpathy, Ph. D.

# ACKNOWLEDGEMENT

I would like to specially acknowledge the support and guidance provided by my committee chair, John Nofsinger. I have a deep sense of gratitude towards him that words can not express. He has helped me learn to be a free thinker and researcher. My dissertation has greatly benefited from the insightful suggestions of my committee members, Richard Sias and Swaminathan Kalpathy. I admire the scholarly aptitude of Richard Sias and thank him for his support. I thank Swaminathan Kalpathy for being a great friend, philosopher and guide. I acknowledge Terrance Odean for providing the discount brokerage house data, which has been used extensively in my dissertation.

The faculty members of the Department of Finance, Insurance and Real Estate at Washington State University have been very helpful and have always looked out for my best interests. I thank them all for the time they invested in me and the many helpful suggestions that helped my shape and sharpen my research. I am especially grateful to Gene Lai and Sandra Boyce for providing their unflinching support throughout my doctoral studies.

I thank my parents, Chander and Shalini Varma, for supporting me throughout my endeavors. Everything I achieve is as much theirs as mine. I treasure the love and encouragement my entire family and close friends have provided over the years. I thank my friend, Urooj Khan for showing me the path to a Ph.D.

INDIVIDUAL INVESTING IN THE FAMILIAR AND

THE UNFAMILIAR

Abstract

By Abhishek Varma, Ph.D.
Washington State University
May 2009

Chair: John R. Nofsinger

Using a discount brokerage house data for the period January 1991 – November 1996, my thesis explores aspects of individual investing behavior in familiar and unfamiliar assets. Behavioral portfolio theory postulates that individuals divide their portfolio into layers of assets with different levels of risk and each of these layers are associated with different aspirations. Individuals are likely to perceive familiar assets to be less risky. I explore familiarity in the following contexts: (1) Individuals frequently repurchasing (purchasing stocks that were previously sold) stocks, and (2) Individuals purchasing their local utility stocks. In the riskier layer of an individual aimed at a shot for the riches, I explore investments in Over-the-counter (OTC) stocks, unfamiliar to most individuals.

Similar to findings in retirement studies focused on employee allocations of their 401(k) plans to their company's stock, I find that repurchases are driven by the interaction of investor preference for familiar stocks and extrapolation of prior roundtrip trade returns. There is no evidence that the repurchasing strategy outperforms a buy and hold strategy. I attribute the sub-

optimality of repurchases to commission costs and under-diversification of portfolios, which are magnified for households repurchasing at higher frequencies.

Individuals in my dataset are nearly four to five times more likely to purchase stocks of their local direct utility as opposed to utility companies operating outside their state of residence. My tests reveal that individuals do not possess superior or private information about their local utilities, nor are they using their local utility stocks as a hedge for possible increase in their utility expenditure. Indeed, individual preference for their local utility stocks seems to be driven by preference for familiar assets, referred to as familiarity bias.

Lastly, I explore commonly held beliefs about individuals investing in OTC stocks, presumably unfamiliar assets. Contrary to popular perceptions associated with gambler or lottery buyers, I find that investors are older, wealthier and more experienced at investing than their counterparts. Individuals investing in OTC stocks display a greater degree of diversification and have large portfolio turnovers.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENT ..... iii

ABSTRACT ..... iv

LIST OF TABLES ..... x

LIST OF FIGURES ..... xii

DEDICATION ..... xiii

CHAPTER 1

OVERVIEW ..... 1

CHAPTER 2

EXTRAPOLATION AND THE FAMILIAR: THE SUB-OPTIMAL REPURCHASING BEHAVIOR OF RETAIL INVESTORS ..... 4

1 Introduction ..... 4

2 Literature Review ..... 9

2.1 Familiarity Bias ..... 9

2.2 Extrapolation Bias ..... 14

3 Data, Methodology and Summary of Repurchasing Activity ..... 17

3.1 Data Description ..... 17

3.2 Measuring Return Performance ..... 20

3.3 Abnormal return performance ..... 22

3.4 Familiarity Bias Statistic ..... 24

3.5 Concentration of Repurchases within few industries and firms ..... 24

4 Testing for Familiarity Bias ..... 26

5 Impact of Extrapolation on Repurchase Decision ............ 29
5.1 Evidence by comparing repurchased and non repurchased stocks among repurchasing households............ 29
5.2 Evidence of investors learning from performance of prior roundtrip trade............ 30
6 Modeling the Decision to Repurchase ............ 32
7 Repurchase: Robustness Check ............ 36
7.1 Are Repurchases mere gambles made in a few select accounts?............ 36
7.2 Do the accounts in which investors conduct repurchases form a significant part of an investor's total wealth and overall common stock trading activity?............ 37
8 Buy and Hold Returns Vs Realized Returns............ 38
9 Sub-optimality of Repurchases ............ 40
9.1 Impact of Commission costs on performance of repurchase roundtrip trades............ 40
9.2 Under-Diversification associated with preference for the familiar............ 41
10 Summary and Conclusion............ 47
CHAPTER 3
INDIVIDUALS AND THEIR LOCAL UTILITY STOCKS: PREFERENCE FOR THE FAMILIAR ............ 50
1. Introduction............ 50
2. Data ............ 55
3. Methodology............ 58
3.1. Return and Performance Measurement............ 58
3.2. Abnormal return performance............ 59
4. Analysis of allocation to local utility stocks at the household level............ 61

6. Explanations for individuals preferring their local utility stocks........................................ 68

6.1. Information Based Explanation ........................................................................................ 68

6.2. Hedging Explanation ........................................................................................................ 71

6.3. Familiarity Bias................................................................................................................ 73

7. Are local utility investors naïve? ........................................................................................ 74

8. Wealth effects and preference for local utility investments................................................. 75

9. Who makes money on their local utility stocks? ................................................................. 76

10. Conclusion ........................................................................................................................ 77

CHAPTER 4

INDIVIDUALS AND OTC (OVER-THE-COUNTER) STOCKS: PREFERENCE FOR THE UNFAMILIAR ..................................................................................................................... 78

1 Introduction.......................................................................................................................... 78

2 Literature review and beliefs about Over-the-Counter (OTC) Stock ................................. 80

2.1 Literature Review.............................................................................................................. 80

2.2 Over-the-Counter (OTC) Stock Description.................................................................... 81

2.3 Why Invest in OTC stocks? .............................................................................................. 84

2.4 OTC stocks and speculative forces .................................................................................. 85

3 Data and Methodology.......................................................................................................... 85

3.1 Data ................................................................................................................................... 85

3.2 Methodology ..................................................................................................................... 87

4 Demographics and OTC Investors........................................................................................ 91

4.1 Univariate tests.................................................................................................................. 91

4.2 Regression results ............................................................................................................ 94

4.3 Robustness (Alternate OTC investor classification) – Regression Analysis ................ 95

5 Portfolio Characteristics and OTC Investors ........................................................................ 96

6 Asset Allocation and OTC Investors ........................................................................................ 99

7 Conditions surrounding trades in OTC stocks ...................................................................... 100

8 Performance of OTC roundtrip trades across OTC quintiles............................................. 101

9 Performance of OTC roundtrip trades across price categories .......................................... 102

10 Conclusion .................................................................................................................................. 103

REFERENCES ............................................................................................................................... 106

LIST OF TABLES


Table 2.1 .......... 111

Table 2.2 .......... 112

Table 2.4 .......... 114

Table 2.5 .......... 115

Table 2.6 .......... 117

Table 2.7 .......... 118

Table 2.8 .......... 120

Table 2.9 .......... 122

Table 2.10 .......... 124

Table 2.11 .......... 125

Table 2.12 .......... 126

Table 2.13 .......... 127

Table 3.1 .......... 130

Table 3.2 .......... 131

Table 3.3 .......... 133

Table 3.4 .......... 135

Table 3.5 .......... 136

Table 3.6 .......... 137

Table 3.7 .......... 138

Table 4.1 .......... 139

Table 4.2 .......... 140

Table 4.3 .......... 142

Table 4.4 .......................................................................................................................... 143

Table 4.5 .......................................................................................................................... 144

Table 4.6 .......................................................................................................................... 145

Table 4.7 .......................................................................................................................... 147

Table 4.8 .......................................................................................................................... 148

Table 4.9 .......................................................................................................................... 149

LIST OF FIGURES

Figure 2.1 .......... 150

Figure 2.2 .......... 151

Figure 3.1(a) .......... 152

Figure 3.1(b) .......... 152

Figure 4.1(a) .......... 153

Figure 4.1(b) .......... 153

Figure 4.1(c) .......... 154

## Dedication

This dissertation is dedicated to my loving parents, Chander and Shalini Varma.

# CHAPTER 1

# OVERVIEW

Using a discount brokerage house trading data for the period January 1991 – November 1996, my thesis explores aspects of individual investing behavior in familiar and unfamiliar stocks. Though the familiar and the unfamiliar are extremes, they are important from the context of individual investing. Shefrin and Statman (2000) develop the behavioral portfolio theory, wherein investors choose portfolios by considering expected wealth, desire for security and potential, aspiration levels, and probabilities of achieving aspiration levels. In contrast, Markowitz's (1952) mean-variance portfolio theory states that investors choose portfolios by considering mean and variance. According to the behavioral portfolio theory (Statman, 2004), investors divide their money into many layers—each of which corresponds to a goal or aspiration. Mean-variance investors have a single attitude toward risk, not a set of attitudes layer by layer. On the other hand, behavioral investors have many attitudes toward risk, so they might be willing to take a lot more risk with some of their money than with other portions.

The study of investing in familiar and unfamiliar assets fits well with the behavioral portfolio theory. Investors feel comfortable investing in stocks they are familiar with. They associate familiarity with lower levels of risk, a belief which might be misplaced in reality. Investments in familiar stocks could be associated with a layer of an individual's portfolio designed for income generation in assets with lower levels of risk. At the opposite end of the spectrum of stocks are unfamiliar stocks, stocks about which investors have no information or familiarity. Investments in unfamiliar stocks could be for speculative reasons and can be associated with a layer of the portfolio designed for a shot at big gains. If I was to look at a pyramid of risk with riskiest point being the tip of the pyramid and the base being the safest part of the portfolio, I would associate investments in stocks with the top half of the pyramid. The lower layers of this top half would associate investment in familiar stocks with lower levels of risk. Investment in penny stocks would form the tip of the pyramid, indicating the riskiest investments alongside lotteries and similar gambling activities. The core of the dissertation is divided into three chapters: Chapter 2, Chapter 3 & Chapter 4.

Individual investors are faced with thousands of stocks from which to select from. Yet they often sell and then repurchase the same stock. Chapter 2 explores the underlying causes and optimality of these repurchases. Similar to Bernartzi's (2001) findings in the context of employee allocations of their 401(k) plans to their company's stock, I find that repurchases are driven by the interaction of investor preference for familiar stocks and extrapolation of prior roundtrip trade returns. Also, there is no evidence that the repurchasing strategy outperforms a buy and hold strategy. I attribute the sub-optimality of repurchases to commission costs and under-diversification of portfolios, which are magnified for households repurchasing at higher frequencies.

In Chapter 3, I document that individuals are nearly four to five times more likely to purchase stocks of their local direct utility as opposed to utility companies operating outside their state of residence. My tests reveal that individuals do not possess superior or private information about their local utilities, nor are they using their local utility stocks as a hedge for possible increase in their utility expenditure. Indeed, individual preference for their local utility stocks seems to be driven by preference for familiar assets, referred to as familiarity bias. In addition, I find that this pervasive behavior can not be attributed to individuals who are less affluent or less sophisticated than their counterparts.

Chapter 4 explores commonly held beliefs about individuals investing in Over-the-counter (OTC) stocks, presumably unfamiliar assets. Contrary to popular perceptions associated with gambler or lottery buyers, I find that investors are older, wealthier and more experienced at investing than their counterparts. Individuals investing in OTC stocks display a greater degree of diversification and have large portfolio turnovers.

# CHAPTER 2

# EXTRAPOLATION AND THE FAMILIAR: THE SUB-OPTIMAL REPURCHASING BEHAVIOR OF RETAIL INVESTORS


**Abstract**

*Individual investors are faced with thousands of stocks from which to select from. Yet they often sell and then repurchase the same stock. Similar to Bernartzi's (2001) findings in the context of employee allocations of their 401(k) plans to their company's stock, I find that repurchases are driven by the interaction of investor preference for familiar stocks and extrapolation of prior roundtrip trade returns. There is no evidence that the repurchasing strategy outperforms a buy and hold strategy. I attribute the sub-optimality of repurchases to commission costs and under-diversification of portfolios, which are magnified for households repurchasing at higher frequencies.*


## 1 Introduction

Today, the US equities market has about 15,000 stocks that are traded on various exchanges. Individuals are faced with information overload in choosing among these stocks because they have limited cognitive ability to assimilate such large amounts of information available to them. This scenario forces them to rely on cognitive shortcuts and heuristics. This study explores how familiarity and extrapolation bias lead people to focus on a few stocks, some of which they repeatedly buy and sell. It also reflects upon the optimality of such behavior.

Bernartzi (2001) finds that about a quarter of the discretionary contributions in large retirement savings plans are invested in company stock and attributed this behavior partly to employees excessively extrapolating past performance (i.e. they take abnormally high past performance as representative of future performance), the effects of which could be magnified by familiarity with the company. Based on a Morningstar.com survey, Bernartzi provided further evidence suggesting that extrapolation behavior is magnified by familiarity. Similar to his study, I explore the interaction of familiarity bias and extrapolation bias in the context of investors repurchasing stocks they previously sold.

It is shown here that individual investors seem to frequently repurchase a stock previously sold. A repurchase could refer to either the purchase of stock that was previously sold or an additional purchase of stock that is already held (like dollar cost averaging). This study focuses solely on the former and not the latter.[1] Analysis of the Barber and Odean (2000) brokerage data of over 60,000 households at a large discount brokerage house between January 1991 and November 1996 demonstrates that this behavior is observed in over 40% of the households analyzed. What is more intriguing is that these 'repurchasing households' account for nearly 80% of stock purchases (848,388 trades) and nearly 90% ($10.55 billion) in dollar values relative to all households (i.e., including those that do not repurchase). Repurchase trades account for 17.1% of the total number and 28.0% in dollar value of all purchase trades made by repurchasing households. The statistics on repurchase transactions are potentially understated due to the unavailability of trading records for these households prior to January, 1991.

The following sample of trading activity in a particular stock extracted from an individual's trading records provides an illustration. An individual purchases 1,000 shares of

[1] It can be argued that dollar cost averaging purchases of a stock are also a ramification of familiarity. Given the few stocks held in the typical brokerage account, the purchase of more of one particular stock forgoes the opportunity to diversify.

Wal-Mart Stores Inc on January 17, 1991 for $32.375 per share and then sells on February 12, 1991 for $36.00 per share, making a net return (including commission costs) of 10.16% on this round-trip transaction in just under one month. Both familiarity and this positive experience may have led to the repurchase of the same stock with a purchase of 1,000 shares on April 13, 1993 for $29.50 per share. Thereafter, this position was partly liquidated with the sale of 500 shares sold on April 5, 1994 at $24.625 per share. Interestingly, the investor did not repurchase Wal-Mart through the end of the data sample. This study posits that the extrapolation of the poor performance in the last round-trip trade, is responsible for lack of any subsequent repurchases of Wal-Mart's stock.

I find no evidence that a repurchasing strategy is superior to a buy and hold strategy. This study argues that repurchasing a stock previously sold is a sub-optimal behavior.[2] Repurchasing involves trading costs, which tend to reduce net returns (Barber and Odean, 2000). In addition, each time an individual investor makes a purchase, he or she has a chance to diversify. Instead, many choose to invest in a stock they held in the past or currently hold in their portfolio. They often confuse familiarity with superiority. Retail investors seem to fixate on stocks that are familiar to them—possibly even having an emotional attachment to stocks in which they held or hold a position. They keep a close watch on these stocks and look to correctly time re-entry after having exited an earlier position in the same stock. This study argues that these stock repurchases are an implication of the familiarity heuristic. An alternative explanation for continually selling and repurchasing a stock is that the investor obtains private information about the stock value in a semi-strong form efficient market. Evidence provided here indicates that this is unlikely

[2] An exception may be the closing out of a short position. Covering short positions are not considered repurchases for this study.

because repurchase trades are not associated with positive alphas and that they seem to occur in very large firms, which are likely to have little information asymmetry.

Barber, Odean and Strahilevitz (2004) in their study of price paths of stocks prior to repurchase find that apart from two other trading patterns, individuals prefer to repurchase stocks that they previously sold for a gain. They argue that this results from a simple form of learning whereby investors repeat actions that previously resulted in pleasure while avoiding actions that previously led to pain. They find that the style adjusted returns earned on repurchase of stocks previously sold for a gain is slightly positive but not reliably different from zero.

This study explores why investors would ever consider repurchasing a stock, irrespective of whether it was previously sold for a gain or loss. I find strong evidence that these transactions are not driven by private or superior information in a semi-strong efficient market, but rather by investors preferring the familiar. This study documents a newer dimension to familiarity associated with repurchases unlike most prior studies that document familiarity bias in the context of geographical location (Cooper and Kaplanis (1994), Huberman (2001) and Grinblatt and Keloharju (2001) among others). Repurchasing stocks are not mere gambles taken by individuals in a few accounts or accounts that form an insignificant portion of their wealth. Investors repurchase stocks that significantly outperform their non-repurchased stocks. I provide further evidence of investors learning from their past actions documented in Barber, Odean and Strahilevitz (2004). Specifically, the average investor makes an additional roundtrip repurchase transactions on a stock only if he or she is able to outperform the market in the prior roundtrip trade.

I also attempt to determine other stock and investors characteristics that might contribute to individuals repurchasing stocks. It seems that investors prefer to repurchase stocks that are in the technology sector, have a large market capitalization, and receive attention based on extremely positive previous day returns. The technology sector preference relates to the general euphoria surrounding technology stocks in the 1990s. Preference for large stocks also captures an additional element of familiarity given the vast amounts of analyst following and news circulation surrounding them. Among various investor characteristics, the most relevant variable relates to client classification assigned by the brokerage house. The results indicate that active traders are more likely to repurchase stocks in comparison to affluent or general traders. This attention bias documented in repurchases is consistent with the findings of Barber and Odean (2006).

To study if the repurchasing strategy is superior to a buy and hold strategy I compared buy and hold returns to realized returns, which were computed with an assumption of investors investing at the risk free rate during the period between the sale and subsequent repurchase of a stock. I find that individuals exhibiting a greater propensity to repurchase seem to have some market timing ability. However, after adjusting for commission cost, though statistically insignificant, buy and hold returns marginally outperform repurchasing return. This study documents that repurchasing stocks frequently is a sub-optimal trading activity in the context of observed performance of repurchase after adjusting for commission costs and diversification. Contrary to Grossman and Stiglitz (1980), I find that individuals repurchase stocks even when the marginal benefit of trading does not exceed the marginal cost. In regard to diversification, individuals with lower familiarity bias are also found to be diversified across more industries and across more stocks with a lower average correlation, indicating an element of skill.

The rest of this study is organized as follows. Section 2 discusses related literature and Section 3 presents data, methodology and summary of repurchases. An exploration of whether repurchases can be attributed to familiarity with a stock is conducted in Section 4. I document the impact of representativeness/extrapolation in Section 5. In Section 6, I attempt to document the impact of factors other than familiarity associated with having held a stock that result in repurchases. Section 7 provides evidence against the arguments that repurchases are gambles and that they probably occur in portfolios that form an insignificant portion of an investor's wealth. Comparison of the performance of a buy and hold strategy against the repurchasing strategy is presented in Section 8. An extensive study of sub-optimality of repurchases is conducted in Section 9. Summary and concluding remarks are presented in Section 10.

## 2 Literature Review

### 2.1 Familiarity Bias

In the words of Huberman (2001), "Familiarity is associated with a general sense of comfort with the known and discomfort with—even distaste for and fear of—the alien and distant." The ramifications of familiarity bias create a tendency for investors to invest in securities familiar to them.

Earlier research indirectly acknowledges the presence of familiarity bias. Merton (1987) develops a model for capital market equilibrium with incomplete information wherein it is assumed that each investor generally knows about only a subset of available securities and this subset differs from one investor to another. They have a familiarity with these stocks in the sense that they only know about the parameters of these stock returns' distribution. The author states that "Recognition of different speeds of information diffusion is particularly important in empirical research where the growth in sophisticated and sensitive techniques to test evermore-refined financial-behavioral patterns severely strains the simple information structure of our asset pricing models." For simplicity, Merton (1987) divided the cost of information transmission into two parts: (1) cost of gathering and processing data (2) cost of transmitting information from one party to another. The results showed that in equilibrium, stocks with a smaller investor base will have lower prices (and higher expected returns).

Heath and Tversky (1991) were the earliest to illustrate the role of familiarity in decision making under uncertainty. In a series of experiments, they showed that "holding judged probability constant people prefer to bet in a context where they consider themselves knowledgeable or competent rather than in a context where they feel ignorant or uninformed." Considering oneself competent does not mean that one is truly competent. They explain the competence hypothesis in terms of asymmetry of credit and blame induced by knowledge or competence—"If the decision maker has limited understanding of the problem at hand, failure will be attributed to ignorance, whereas success is likely to be attributed to chance. In contrast, if the decision maker is an "expert," success is attributable to knowledge, whereas failure can sometimes be attributed to "chance." This study concluded that their results "might explain why investors are sometimes willing to forego the advantage of diversification and concentrate on a small number of companies (Blume, Crockett and Friend, 1974) with which they are presumably familiar." Therefore, the ramification of familiarity is that investors primarily focus on familiar stocks and fail to fully diversify.

In the past, familiarity has been mostly explored in the context of: geographical proximity, cultural proximity, professional proximity, language preferences, or political setup. The following sub-sections provide evidence of familiarity bias from prior research.

### 2.1.1 Evidence from literature on Home Bias

Grubel (1968), Levy and Sarnat (1970), Solnik (1974) and many other studies recommend international diversification for the potential gains it offers to investors. Lewis (1999) shows that a portfolio with a 100 percent share in the S&P 500 is dominated by all portfolios with a minimum foreign share of about 39 percent, corresponding to the minimum variance point on the efficient frontier. However, the literature is replete with evidence of investors being under diversified internationally—representing one of the biggest unresolved puzzles in international finance.

For example, French and Poterba (1991) estimate the domestic ownership of shares in the worlds' five largest stock markets at: US (92.2%), Japan (95.7%), UK (92%), Germany (79%) and France (89.4%). Tax rules, transaction costs and explicit limits on cross-border investment among these developed countries fail to explain this home bias. They attribute the lack of diversification to investor choice rather than institutional constraints. They find that investors seem to expect the returns in domestic markets to several hundred basis points higher in comparison to expectations of other markets. The fear of the unknown, or unfamiliar, seems to be an underlying reason for higher expected returns, which in turn leads to home bias among investors.

Cooper and Kaplanis (1994) find that the home bias cannot be explained by inflation hedging or directly observable costs of international investments unless investors have very low levels of risk aversion. Tesar and Werner (1995) provide evidence of home bias and the inability of transaction costs to explain this bias. Geographical proximity is found to be an important ingredient of the international portfolio allocation decision in their study.

Previous studies mostly studied home bias from the perspective of US investors. Chan, Covrig and Ng (2005) examined mutual funds from 26 developed and undeveloped countries to determine factors affecting asset allocation worldwide. They divided home bias into two components: (1) Domestic bias – Extent to which mutual fund investors overweight home markets in their mutual fund holdings, and (2) Foreign bias – Extent to which investors underweight or overweight foreign markets. This division enabled them not only to see how each bias varies across countries, but also to examine whether the investment barriers have similar or different impacts on the two biases. Their results showed that stock market development and familiarity variables play an important role in domestic bias. These two variables also exhibit significant but asymmetric, effects on foreign bias. Other factors such as economic development, capital controls, and withholding tax variables have significant effects only on foreign bias.

### 2.1.2 Evidence from literature on within country Geographical Bias

Similar to the home bias, people have an affinity for investing in stocks in the local geographical area. Huberman (2001) considers the geographical distribution of the shareholders of the seven U.S. Regional Bell Operating Companies (RBOCs) at the end of 1996 and finds that in most states, more money is invested, per investor, in the local RBOC than in any other RBOC. He finds no support for the argument that a customer of an RBOC may over-invest in its stock as a hedge against unexpected increase in the price of its services. This familiarity bias contradicts traditional portfolio theory, which implies that investors should diversify and invest less in the RBOC serving him than in those other parts of the country since the fortunes of the RBOCs vary with the economic tides in their home areas. Grinblatt and Keloharju (2001) in their study of the Finnish stock market find that investors simultaneously exhibit a preference for nearby firms and for same-language and same-culture firms. Massa and Simonov (2006) studied a unique data of Swedish investors and find that investors do not hedge, but instead, invest in stocks closely related to their non-financial income. They state that stocks holdings were driven by geographical and professional familiarity and that this familiarity is not a behavioral bias, but is information driven. Coval and Moskowitz (1999) found that U.S. investment managers, in a setting of a single currency and relatively little geographical variation in regulation, taxation, political risk, language, and culture, prefer to hold companies headquartered close to them. Their results suggest an information based explanation for local equity preference since the firms preferred by the investment managers tend to be small and highly levered, and they tend to produce goods not traded internationally (ie. firms with greater information asymmetry).

Prior studies have provided alternative explanations to the home bias mentioned earlier. Serrat (1997) has shown that in an international exchange economy with two agents the home bias puzzle can be attributed to non-tradability of some goods affecting the marginal utility of tradable goods. Stulz (1981) provides evidence on how restrictions on international capital flows could lead to the preference for domestic assets. However, the same arguments cannot be used to explain the impact of familiarity in domestic investments. Thus there remains strong evidence of investors preferring the familiar.

### 2.1.3 Evidence from literature on Retirement Savings Behavior

Studies have shown that investors tend to invest a large portion of their retirement savings in their employer's stock. Benartzi (2001) finds that employee's discretionary contribution to company stock varies according to the past stock performance but does not predict future performance, which is consistent with the excessive extrapolation hypothesis. This evidence does not support the hypothesis that employees possess private information about their company's stock. Also, employees could be viewing employers' contribution in the form of company stock as an endorsement or implicit advice to invest their discretionary component in company stock (Benartzi, 2001). The findings generally support the notion that investors prefer stocks familiar to them and also imply that this bias may not operate in isolation. The positive correlation between past return and allocation of discretionary contribution in company stock provides evidence of representativeness interacting with familiarity bias.

Investing ones financial and human capital in the same company is unwise. If for some reason a company witnesses turmoil, not only do its' employees suffer financial losses in their retirement portfolio but could also lose their jobs. The Enron and WorldCom fraud in 2002 transformed many older employees from paper millionaires into poppers. However, investors do not seem to be learning from the past and still prefer the familiar. A survey by Boston Research Group (2002) finds that the Enron fiasco has had little effect on 401(k) participants holding of company stock. The survey finds that 401(k) participants who invest in their company stock have 30% of their 401(k) assets in company stock, with one-third of participants holding more than 50% of their assets in company stock. At present, there is an intense debate between policy makers, academicians, and the financial press, on whether a limit on employee allocation of 401(k) contributions to company stock should be put in place. Benartzi and Thaler (2007) question why the congress permits the use of company stock in 401(k) plans.

## 2.2 Extrapolation Bias

The extrapolation bias draws from the seminal work of Tversky and Kahneman (1974) on representativeness bias. Representativeness is one of the most important heuristics used by individuals to reduce the complex task of assessing probabilities and predicting values to simpler judgmental operations and can be considered a bias of stereotypes (Shefrin, 2005). This heuristic leads people to form probability judgments that systematically violate Bayes's rule (Kahneman and Tversky, 1972 and Tversky and Kahneman, 1974).

A naive investor who has no prior knowledge of the underlying return distribution of the stock would extrapolate the extremeness of large occurrences of positive returns (strength of the evidence) relative to the small sample size (the weight of evidence) and conclude that the expected return is likely to be positive as well. Griffin and Tversky (1992) find that people are highly sensitive to variations in the extremeness or strength of evidence and not sufficiently sensitive to variations in its weight, credence or predictive validity. Motivated by this idea, Barberis et al. (1998) formed a model for investor sentiment where stock prices underreact to news single events such as an earning announcement, but overreact to a series of good or bad news announcements.

Investors impacted by the extrapolation bias consider recent past returns to be representative of what they can expect in the future. Because of this cognitive error, investors might buy recent past winner stocks and sell (or avoid) past losers. Based on nearly 38,000 forecasts of stock prices, De Bondt (1993) finds that individual investors seem to be trend followers, predicting price trend continuation. Similarly, Grinblatt and Keloharju (2000) find that foreign investors in Finland both buy past winners and sell past losers, thus showing an extrapolation behavior in both purchases and sales. Chen et al. (2007) conclude that Chinese investors focus on the most recent performance of the stocks they purchase—they buy past winners. This extrapolation trading behavior is consistent with investor beliefs about expected returns. Bernartzi (2001) asked Morningstar.com subscribers to rate the performance of company stock over the last five years and the next five years. Despite the unpredictability of returns, the respondents' ratings were positively correlated ($\rho = 0.52$), which is consistent with the extrapolation hypothesis. Vissing-Jorgensen (2003), in an extensive survey of 1000 households holding financial assets in excess of $10,000, finds that an individual's belief about future stock market returns depend on his or her past (self-reported) portfolio returns.

Value strategies call for buying stocks with low price relative to earnings, dividends, book assets or other measures of fundamental value. Solt and Statman (1989) document that the higher the growth opportunities for a company, the lower the risk-adjusted return that its stock provides to shareholders in the subsequent period. Lakonishok, Shleifer and Vishny (1994) provide evidence that value strategies yield higher returns than naïve strategies pursued by other investors because they exploit sub-optimal behavior of investors and not because they are fundamentally riskier. Amongst other explanations, they state that naïve strategies might involve extrapolation of past earning too far into the future, a prediction of the extrapolation hypothesis.

Research studying the sub-optimal behavior of employees concentrating a large part of their 401(k) assets in the company stock also reports that this behavior is exacerbated when the company's stock has recently performed well (Benartzi, 2001). Bernartzi formed five portfolios on the basis of past buy and hold return and examined the subsequent discretionary contributions to company stock. When portfolios were formed on the basis of 10 year returns, the lowest-return portfolio has 10.37 percent allocated to company stock versus 39.70 percent for the highest-return portfolios.

In addition, even the choices of funds available in the pension plan determined by the plan administrator appear to be highly influenced by past returns. Elton et al. (2007) study the performance of mutual funds offered by 401(k) plans. They find that the decision to add or drop funds seems to be made, at least in part, on the basis of past performance. Interestingly, the funds that were added to these plans performed no better than the funds that were dropped.

## 3 Data, Methodology and Summary of Repurchasing Activity

### 3.1 Data Description

The data covers the investment and trading activity of households at a large discount brokerage house during the period January 1991 to November 1996. There are 77,995 households in the dataset and each household has an average of 2.02 accounts with the brokerage house. This study focuses only on the common stock trading activity of these households. During the sample period 80.7% percent of households (62,942 households) traded in common stocks. There are 1,969,747 trades in the dataset including 11,318 trades that are reversal entries. Correcting for these reversals and deletion of 47 zero-dollar value trades, there are 1,947,298 trades left with slightly more purchases (1,071,182) than sales (876,116).

A brief description of the price, size and commission costs of all trades is provided in Table 2.1. Panels A and B indicate that not all the households make both purchases and sales. The number of households that make purchases (55,902) is less than the number of households that make sales (56,997). The total value of all trades equals $23.87 billion with purchases and sales, both valued at $11.94 billion. The households traded 15,493 stocks during the 71-month period, out of which 10,186 stocks were traded on the NYSE, AMEX or NASDAQ, accounting for 83% of all stocks (12,290) listed on these three exchanges during the same period. For the remaining 5,307 stocks there is no information in the CRSP database. Panel C of Table 2.1 shows that 75 percent of non-listed stocks are priced below $3.12. This leads to the conclusion that most of these stocks are penny stocks and they have been excluded from further analysis as they account for .92% (refer Table 2.2, Panel B) of total purchases in value terms.

A "repurchase" refers to buying a stock that was previously sold. Households that make at least one repurchase are termed as a "repurchasing household." Repurchasing households represent 41.09% (22,971 households) of all households that made purchases. Interestingly, these households represent nearly 80% (848,388 purchase trades) of the number and nearly 90% ($10.55 billion) of the dollar values, of all purchase trades. Repurchasing stocks previously sold seems to be a more pervasive trading activity among the households who dominate the trading activity among all households.[3] Table 2.2 provides a description of trading activity based on the NYSE/AMEX/NASDAQ[4] size deciles formed on 29 November, 1996.[5] Panels B and C show that the number of stocks traded (purchase or sales) seems to be evenly distributed amongst various size deciles except for the lowest decile that accounts for the largest number (1,402) of stocks. However, the majority of the trading is concentrated in the large market capitalization stocks. In terms of dollar value of purchases (sales), the top two deciles representing market capitalizations in excess of $534 million[6] account for 75% (77%) of all purchases (sales). These deciles also represent 62% (64%) of the total number of purchase (sale) trades.

Repurchase trades account for 17.07% (144,829 purchases) of the total number and 28.03% ($2.95 billion) of the dollar value of all purchase trades by the repurchasing households.

[3] Repurchasing households account for 78% of number and 83% of value of all trades (ie. purchases and sales). Thus the figures in Table 2.1 largely depict the trades made by repurchasing households.

[4] There are 4,058 stocks (ie. 30% of all stocks traded) for which no data is available in CRSP and could not be allocated to any size decile. However this does not have any significant impact on the data descriptions as trading in these stocks in terms of percentage of dollar value of purchases (sales) and percentage of number of trades accounts is 2.41% (2.16%) and 3.54% (3.41%) respectively. As stated before, the market capitalization of these stocks is unlikely to exceed capitalization of the lowest market decile stock in the CRSP database.

[5] The last date on which any households traded stocks is 29 November, 2006. NYSE/AMEX/NASDAQ deciles are formed by extracting data on market capitalization of all stocks with a non-zero market capitalization and cut-offs for these deciles are used to segregate trades into various size deciles.

[6] The market capitalization range for each decile is provided in Table 2.2, Panel A

In Table 2.3, repurchasing households are categorized based on the number of repurchases made to investigate if there are any cross sectional differences within these categories.[7] The majority of the households make only one repurchase and the numbers of households decrease with higher repurchases. All figures except for the number of households are computed by averaging across all households in a particular category. Repurchases as a percentage of all purchase trades stood at 17% for the lowest repurchasing category. This figure increases monotonically to a high of 36% for the highest repurchase category. In terms of dollar values of repurchases, I observe the same pattern. Category A shows a low of 20% while Category K shows a high of 42%. Thus, households that repurchase more frequently, also repurchase a greater amount relative to both the number of total purchases and the value of total purchases. The number of stocks repurchased shows that on average, the households repurchasing more frequently do not seem to be concentrating their repurchase in a single stock. The category K shows that households that repurchase between 51 to 100 times on average repurchase 32 stocks. The monotonic trend in the mean monthly turnover for each category indicates that households that repurchase more frequently also experience greater turnover.[8] The lowest repurchase category is likely to turnover 84 percent of its portfolio annually while higher repurchaser, like category G, are likely to turnover 180 percent of their portfolio annually. The number of stocks traded is calculated over 71 months and cannot by itself be used to indicate portfolios diversification. Even a figure of 44 stocks for category H households indicates insufficient diversification, given that these households have an extremely high turnover of 216 percent annually. The last column in Table

[7] Number of times a household repurchases is not specific to a particular stock. For example, if a household repurchases the stock of Microsoft Corp twice and Boeing's stock once, the total number of repurchases is equal to three.
[8] The monthly purchase and sales turnover are calculated accordingly to the methodology followed in Barber and Odean (2000). The average of the purchase and sales turnover is the mean monthly turnover.

2.3 shows that stocks repurchased as a proportion of stocks purchased tends to be higher for households that repurchase more, with a high of 43% for the highest repurchase category.

Table 2.4, Panel A, illustrates the time difference (in days) between the sale and repurchase of a stock across repurchase categories. Overall, the time difference within each category is positively skewed. The households that repurchase more frequently seem to repurchase at shorter intervals of time, which is intuitive as they are more likely to be active traders. Panel B shows all the repurchase trades are categorized based on the number of days between the sale of a stock and its subsequent repurchase. The tables show that 37% of the repurchase trades are made within 1 month of the last sale of the stock and 72% are made within 6 months. This indicates that investors seem to be keeping a close watch on their stocks and within a short period repurchase it for better or for worse.

### 3.2 Measuring Return Performance

This study focuses on the performance of repurchase trades rather than the performance of the whole portfolio held by repurchasing households. Analysis of each round trip transaction helps in testing whether an information-based hypothesis that could explain repurchasing behavior and illustrates the conditions under which repurchases are likely to occur.

The household trading records indicate the prices at which trades take place along with the commission costs. The return computations are made in three steps. Firstly, the gross and net returns for a roundtrip transaction j on security i purchased on day 1 and sold on day T is calculated as:

$$R_j^{Gr} = \left[ \prod_{t=1}^{T} (1 + R_{i,t}^{Gr}) \right] - 1 \quad and \quad R_j^{Net} = \left[ \prod_{t=1}^{T} (1 + R_{i,t}^{Net}) \right] - 1,$$

where the daily returns for all days except t=1 and t=T are obtained from CRSP. The gross and net returns on day of purchase ($R_{i,1}$) and day of sale ($R_{i,T}$) are calculated as:

$$R_{i,1}^{Gr} = P_{i,1} / P_{i,b} \quad and \quad R_{i,1}^{Net} = P_{i,1} / P_{i,b}(1 + c_{i,b})$$
$$R_{i,T}^{Gr} = P_{i,s} / P_{i,T} \quad and \quad R_{i,T}^{Net} = P_{i,s}(1 - c_{i,s}) / P_{i.T} \ ,$$

where $P_{i,1}$ is the closing price on day of purchase, $P_{i,b}$ is the purchase price, $P_{i,s}$ is the sale price, $P_{i,T}$ is the opening price on day of sale, $c_{i,b}$ is the commission cost as a percentage of purchase and $c_{i,s}$ is the commission cost as a percentage of sale. Secondly, the gross ($R_h^{Gr}$) and net ($R_h^{Net}$) returns for all repurchases made by a household h during the period January, 1991 to November 1996 is calculated as

$$R_h^{Gr} = \sum_{j=1}^{n_j} w_j R_j^{Gr} \quad and \quad R_h^{Net} = \sum_{j=1}^{n_j} w_j R_j^{Net}$$

where $R^{Gr}{}_j$ refers to the gross return and $R^{Net}{}_j$ refers to the net return on round trip transaction j, and $w_j$ refers to the dollar value of transaction j scaled by the dollar value of all repurchases made by household h. In the last step gross ($R_k^{Gr}$) and net ($R_k^{Net}$) returns for all repurchasing households assigned to a particular category are calculated as:

$$R_k^{Gr} = \frac{1}{n_k} \sum_{h=1}^{n_k} R_h^{gr} \quad and \quad R_k^{Net} = \frac{1}{n_k} \sum_{h=1}^{n_k} R_h^{net}$$

where $n_k$ is the number of households in a particular category k. The methodology for calculating value weighted returns for each household and the average returns for households is somewhat similar to Barber and Odean (2000). Barber, Odean and Strahilevitz (2004) in their study of repurchases measure performance of portfolios of repurchased stocks previously sold in the previous year over a holding period of 1 year. However, this study substantially differs from all previous studies by measuring the actual value weighted returns of repurchase roundtrip trades.

### 3.3 Abnormal return performance

In this study, six measures of abnormal returns for roundtrip trades are considered. The first measure is the market adjusted return, obtained by subtracting the market return from the roundtrip trade return. The CRSP value weighted NYSE/AMEX/NASDAQ index is used to proxy for market returns and the return on this index is measured over the duration of each roundtrip trade.

The second measure is the industry adjusted return, obtained by subtracting the industry return from the roundtrip trade return. Industry return is calculated as value weighted return of all firms in the same industry excluding the repurchased stock and is measured over the duration of each roundtrip trade. A large portion of repurchases are concentrated in a few industries like the technology sector that experienced phenomenal growth and this measure should reflect on performance relative to other stocks in the same industry. The 49 industry classifications are obtained form Kenneth French’s website.

The third measure used is Jensen’s alpha (α):

$$\alpha_{j,t} = (R_{j,t} - R_{f,t}) - \hat{\beta}_{j,t}(R_{m,t} - R_{f,t})$$

where $\alpha_{i,t}$ is the market model abnormal return for a roundtrip transaction j that is conducted over t calendar days, $R_{i,t}$ is gross or net return on the roundtrip transaction, $R_{f,t}$ is the one month Treasury bill rate scaled to time t. $R_{m,t}$ is the market return on the CRSP value weighted NYSE/AMEX/NASDAQ index measured over time duration of t calendar days, and $\beta_{i,t}$ is the beta in the CAPM. To adjust for any impact of non-synchronous trading the beta is estimated using the Scholes and Williams (1977) procedure over a time window of 250 trading days prior to each transaction.

The fourth measure uses the Fama and French (1993) 3-factor abnormal return and is calculated as:

$$\alpha_{j,t} = (R_{j,t} - R_{f,t}) - \hat{\beta}_1(R_{m,t} - R_{f,t}) - \hat{\beta}_2(SMB) - \hat{\beta}_3(HML)$$

where SMB is a size factor, HML is a value factor and the description for remaining terms is the same as in the CAPM described above.

The fifth measure used is the obtained from the Carhart (1997) 4-factor model and is calculated as:

$$\alpha_{j,t} = (R_{j,t} - Rf) - \hat{\beta}_1(R_{m,t} - R_{f,t}) - \hat{\beta}_2(SMB) - \hat{\beta}_3(HML) - \hat{\beta}_4(WML)$$

where the WML is a momentum factor and the description for remaining terms is the same as in the Fama and French 3-factor model.

The last measure uses a five factor model that adds an industry factor to the Carhart (1997) 4-factor model and is calculated as:

$$\alpha_{j,t} = (R_{j,t} - Rf) - \hat{\beta}_1(R_{m,t} - R_{f,t}) - \hat{\beta}_2(SMB) - \hat{\beta}_3(HML) - \hat{\beta}_4(WML) - \hat{\beta}_5(IND)$$

where IND is the industry return calculated as a value weighted return of all firms in the same industry excluding the repurchased stock. The IND factor is used to capture any industry effects that might be driving returns. The coefficients for all the factors in the Fama and French (1993), Carhart (1997) and the five factor model are measured over a time window of 250 trading days prior to the repurchase trade. The data for $R_f$ (1 month Treasury Bill rate), SMB, HML and WML factors is obtained from Kenneth R. French's online data library.

### 3.4 Familiarity Bias Statistic

The number of repurchases could be used to proxy for familiarity bias. Although this measure provides a good description of the data, it is a crude measure at best. An individual making 1 repurchase trade among 2 purchase trades during the entire sample period is likely to have a greater familiarity bias in comparison to an individual who makes 5 repurchase trades among 30 total purchase trades. In this section a better measure that proxies for a potential familiarity bias is developed and is calculated as:

$$FB\ Statistic_h = \frac{Number\ of\ Repurchases}{Number\ of\ Purchases}$$

where FB $\text{Statistic}_h$ is the proportion of purchase trades made by household h that consists of repurchases. In the following analysis all households are categorized into quintiles based on the FB Statistic, referred to as Familiarity Quintiles. The familiarity quintiles help to test if there are any differences between the households with high and low degrees of familiarity bias.

### 3.5 Concentration of Repurchases within few industries and firms

This section explores the characteristics of stocks in which investors concentrate their repurchasing activity. Table 2.5 presents a detailed distribution of the value of repurchase trades across 27 industry segments[9] and the top three firms with the greatest concentration of repurchase within each industry segment. The high concentration of repurchases in the technology sector stands out in this table, which is expected given the euphoria surrounding technology stocks in the 1990s. The computer hardware, electronic equipment and computer

[9] The definition of industry segments is obtained from Kenneth R. French's online data library. The SIC codes for each stocks are obtained from CRSP. There are a total of 49 industrial segments among which the repurchase trades are classified. Information relating to only 27 sectors is presented as these segments account for most of the repurchases (in specific 93% of total value of repurchases).

software industry segments account for 36% of the total repurchase trades. Also interesting is the high concentration of repurchasing activity in a few firms within most industrial segments. The mean and median percentage of the value of repurchases in the top three stocks within each sector, measured across all major industry segments is 52.68% and 47% respectively. Within each of their respective industry segments, the percentage concentration for Wal-Mart, Intel, General Motors, Motorola, and Pepsico is 21%, 29%, 32%, 46% and 54% respectively. Boeing Corp and Phillip Morris Corp have a very large concentration of 87% and 95% respectively, which can partly be attributed to the existence of a monopoly in their respective industries. The top 3 firm in the utilities sector enjoyed a total concentration of 20%, the least among all industry segments and this can be attributed to the large number of regional utilities known only to local investors.

Individuals seem to concentrate their repurchases in a few large stocks. A casual glance through the list of the 81 firms indicates that even a person with little or no knowledge of financial markets should have some awareness about these firms. In Table 2.2, the highest NYSE/AMEX/NASDAQ capitalization decile accounts for 53.79% of total repurchases in 812 stocks. However, as little as 81 firms presented in Table 2.5, account for 42% of the total repurchasing activity. In other words, 10% of stocks repurchased within the highest capitalization decile makeup for nearly 80% of the repurchasing activity. Given that most of these firms are highly visible with large amounts of news circulation, experience frequent tracking by analysts, and produce products often consumed by investors, it should be no surprise to see investors tilt their repurchases to these stocks as they probably feel more comfortable with these familiar stocks.

Given that a sizable portion of the repurchases occur in large capitalization technology stocks, the average beta of stocks repurchased is likely to be greater than 1. A look at the average betas across various familiarity quintiles, presented in Table 2.6 illustrates if there are any differences in stock characteristics that could explain differences in performance. The average beta of the stocks ranges from 1.36 for the lowest familiarity quintile to 1.47 for the highest quintile. It seems that the higher familiarity quintile investors take on greater market risk, but this could be simply a result of more roundtrip returns being made by investors in high beta technology stocks. Though the F-Test indicates that differences in the betas across the familiarity quintiles are statistically significant, the differences are economically small. The evidence indicates a high degree of homogeneity in the stocks repurchased by households. Thus, differences, if any, in performance of repurchase roundtrip across households can not be attributed to differences in systematic risk among stocks.

## 4 Testing for Familiarity Bias

Repurchases by retail investors could be attributed either to their familiarity with stocks associated with having held a position in them, or to the possession of superior or private information in a semi-strong form efficient market. The former is referred to as the familiarity hypothesis and the latter is referred to as the information hypothesis. Familiarity bias gives investors an illusion of superior knowledge whereas the information hypothesis refers to the actual possession of superior knowledge. The previous sections shows that investors tend to concentrate their repurchases in a few large, highly visible firms. It seems unlikely that investors have private information about these stocks. In this study, the performance of roundtrip repurchase trades is analyzed to determine which of the above two hypotheses explains the repurchasing behavior.

If the information hypothesis were to hold true then an investor should earn positive abnormal returns gross of trading costs. However, if the familiarity heuristic explains the repurchasing behavior, then insignificant or negative abnormal returns should be observed on average across households. Insignificant abnormal returns should be observed if investors were to randomly invest in a security, which shows that familiarity might be unimportant in the context of investment performance. Negative abnormal returns reflect the sub-optimality of repurchasing familiar stocks. The test of information versus familiarity bias can be stated as:

$$\text{H0: Information hypothesis} \Rightarrow \text{Gross Abnormal Returns} > 0$$

$$\text{H1: Familiarity hypothesis} \Rightarrow \text{Gross Abnormal Returns} \leq 0$$

Table 2.7 presents the gross return performance of repurchase roundtrip trades, for the average household in each familiarity quintile. The nominal gross returns range from 9.27% for the highest familiarity quintile to 13.38% for the lowest familiarity quintile. The positive returns are not surprising given that 36% of all repurchases (in value terms) are concentrated in the technology sector, which experienced phenomenal growth during the 1990s. The NASDAQ composite grew at 23% annually during the period 1991-1996.

The market adjusted gross returns presented in Panel A are negative for the lowest three familiarity quintiles and are positive for the highest two familiarity quintiles. Evidence presented in section 5 indicates that investors tend to extrapolate returns earned in the prior roundtrips trades. If a household makes repetitive repurchase roundtrip trades in the same stock, then the returns on these trades, on average, are likely to be positive. Thus, positive market adjusted gross returns do not necessarily mean that these households possess superior information. They could simply be extrapolating excessively by conducting roundtrips at a high frequency in the stocks that have been doing well. Households in the top familiarity quintile, on average, make approximately two roundtrip repurchase trades in every stock they repurchase, whereas households in the lowest familiarity quintile makes only a single roundtrip repurchase trade (refer to Table 2.7, Panel C). The performance of certain sectors of the economy could also be contributing observed market adjusted returns (eg. performance of technology stocks). Analysis of industry adjusted returns shows roundtrip trades on repurchases stocks significantly underperforms peer industry stocks. The highest industry adjusted return is -9.17% for the fourth familiarity quintile and the lowest is -16.04% for the lowest familiarity quintile.

The abnormal gross returns measured using various factor models (refer Panel B, Table 2.7) is negative for all familiarity quintiles, indicating adverse effects of the familiarity bias. The F-Statistic indicates that for all measures of abnormal returns, the means across all familiarity quintiles are jointly significant. For the top two familiarity quintiles the abnormal returns based on the Fama and French (1993) and the Carhart (1997) models is negative but insignificantly different from zero. The explanation of these quintiles outperforming the lowest three quintiles relates to the evidence of greater number of roundtrips in the top two familiarity quintiles discussed in the previous paragraph. With the addition of the industry factor to the Carhart (1997) 4-factor model, the abnormal gross returns become significantly negative for all familiarity quintiles. Overall, irrespective of the familiarity quintile, roundtrip repurchase trades fail to generate positive gross abnormal returns. The evidence presented largely supports the notion that these investors do not possess any superior information and that their repurchase decision is partly driven by their familiarity of the stocks.

## 5 Impact of Extrapolation on Repurchase Decision

### 5.1 Evidence by comparing repurchased and non repurchased stocks among repurchasing households

Although repurchasing stocks seems to be a pervasive trading activity among households that trade frequently, not all stocks are repurchased by these households. Clearly, familiarity with a stock is not the only reason for repurchasing a stock. Exploration of the roundtrip transaction return difference between stocks that are repurchased versus stocks that are not repurchased provides an insight into conditions under which households might decide to repurchase certain stocks.

Table 2.8 explores the differences between non-repurchased and repurchased stocks, among the sample of repurchasing households. The sample analyzed consists only of repurchasing households for which data is available on both roundtrip returns in non-repurchased stocks and repurchase stocks. In Panel A, the number of roundtrips in repurchased and non-repurchased stocks is presented. The average number of roundtrip trades by the lowest familiarity quintile households in non-repurchased and repurchased stock, was 18 and 2 respectively. The proportion of roundtrip trades in non-repurchased to repurchased stocks monotonically decreases from the lowest to the highest familiarity quintile, which is expected based on the calculation of the familiarity bias statistic for each household. The duration of the average roundtrip trades measured in calendar days is significantly lower for repurchased stocks in comparison to non-repurchased stocks, indicating that investors seems to be reluctant to realize losses. This is consistent with the evidence provided by Odean (1998) in support of the disposition effect, the tendency of investors to hold losing investments too long and sell winning investments too soon.

Panel A shows that the gross return for repurchased stocks exceeds those of non-repurchased stocks by a highly significant 10%, on average, across all familiarity quintiles. The differences in performance between non-repurchased and repurchased stocks are further magnified when market adjusted returns and industry adjusted returns are presented in Panel B. The difference in gross market adjusted returns range between -10.97% to -14.15%. The differences based on various other abnormal return calculations (refer Panel C) are at least -10.56%. Expectedly, the results are similar based on net returns as well but have not been presented to save space. Irrespective of the degree of familiarity, investors are more willing to bet on stocks that performed comparatively well in prior roundtrip trades.

Extrapolation of past returns by individual investors has been observed in retirement savings allocation (Benartzi, 2001). The evidence presented in this study shows that while making their decision to repurchase stocks, retail investors could actually be extrapolating returns of their "winning" trades. In other words they want to relive the moment.

### 5.2 Evidence of investors learning from performance of prior roundtrip trade

The analysis in this section provides further evidence supporting extrapolation of returns by individual investors. It shows how investors learn from their prior roundtrip trades in each stock. The previous section involved analysis of repurchase trades aggregated at the household level, ignoring the identity of stocks repurchased. In comparison, this section involves analysis of roundtrip trades by each household in each stock and covers the entire dataset of 62,942 households that traded at the brokerage house. For analysis, a roundtrip serial number is assigned to each roundtrip trade conducted by each household in each stock. In the next step, all roundtrip trades grouped under each serial number are placed in either of the following two categories: (1)

No further roundtrip trades (2) More roundtrip trades. “No further roundtrip trade” consists of stocks in which a household does not conduct any further repurchases. Finally, the average return for all roundtrips with the $n^{th}$ roundtrip is calculated. The following example should provide more clarity: Household A conducts 2 roundtrip trades in Intel and 1 roundtrip trade in Boeing Corp, while household B conducts only 1 roundtrip trade in Intel. The roundtrip trades by household A in Intel are assigned to the $1^{st}$ and $2^{nd}$ roundtrip category. Similarly, roundtrip trades by household A in Boeing and the roundtrip trade by household B in Intel are each assigned to the $1^{st}$ roundtrip category. Thereafter, for the $1^{st}$ roundtrip, the roundtrip trades by households A & B in Boeing and Intel respectively are placed in the “No further roundtrip trade” category while the trade $1^{st}$ roundtrip made by household A in Intel is placed in the “More roundtrip trade” category. Similar categorizations are made for subsequent roundtrips.

The advantage of this methodology is that for each roundtrip trade serial number, the average return on a stock which is repurchased by a household can be compared to a stock for which no further repurchases are made. The results showed that the returns on the trades that led to further repurchases significantly exceeded the return on trades that resulted in no further repurchases. A snapshot of some of the interesting results is presented in Figure 2.1, depicting the performance of market adjusted net returns for each roundtrip repurchase trade. The returns beyond 20 repurchases made in a stock by a household have been ignored as there are very few observations,[10] leading to averages that are insignificantly different from zero. It seems that whenever the market adjusted net returns on a stock turns negative, individuals abstain from making further roundtrip trades in the same stock. However if the returns are positive, they continue to make further roundtrip repurchase trades. For roundtrip number 1, the “No further Roundtrip” and “More Roundtrip” the net market adjusted returns are -3.43% and 7.75%

[10] Less than .01% of total roundtrip repurchase trades.

respectively. It seems like a positive market adjusted net return over a roundtrip trade in a stock is a pre-condition to conducting further repurchases. These conclusions are based on average behavior only. There are quite a few households who continue to make roundtrip trades even after having been burnt in a previous repurchase, but they too generally have earned substantial positive returns in most roundtrips prior to that last roundtrip trade. The evidence presented illustrates how investors learn from their prior performance. Barber, Odean and Strahilevitz (2004), using a measure of the number of winning and losing stocks, found that individual's tend to repurchase stocks previously sold for a gain. The results in this study provide a more detailed description of individuals learning from prior trades by closely tracking the trading activity of individuals.

## 6 Modeling the Decision to Repurchase

In this section an attempt is made to model the impact of stock and investor characteristics on repurchase decisions. The models only study the impact of factors other than familiarity bias associated with repurchases. Households that never make repurchases are assumed to have no familiarity bias and are not included in the analysis. Also, since these non-repurchasing households account for very low levels of trading it makes sense to exclude them from this analysis as they could potentially be trading actively at some other brokerage house. The sample used for these models is the same as the one in Section 5.1, consisting of roundtrip trades made on repurchased and non-repurchased stocks.

The decision to repurchase is modeled using logistic regressions where the independent variable is a dichotomous variable that takes a value 1 if a repurchase trade is conducted on a stock and value 0 if the stock is never repurchase by a household. Stock characteristics and investor characteristics are the dependents variables. The stock characteristics variables are: (1) Positive Net Market Adjusted Returns, (2) Technology sector, (3) Large Stock, and (4) Extreme Positive Previous Day Return. Positive Net Market Adjusted Returns is a dummy variable that takes the value 1 for stock that earned positive market adjusted returns in its last roundtrip repurchase and 0 otherwise. The variable is included based on the evidence presented in Section 5.2 wherein investors seem to learn from the market adjusted returns earned in the previous roundtrip trade. Technology Sector is a dummy variable capturing the impact of investors' preference for technology stocks in the 1990s and takes the value 1 for a stock in the technology sector and 0 otherwise. Large Stock is a dummy variable indicating the investors' preferences for large market capitalization stocks and takes the value 1 if the stock lies in the top market capitalization decile of all NYSE/AMEX/NASDAQ stocks and 0 otherwise. Investor preference for technology sector and large market capitalization stocks captures an additional element of familiarity bias apart from repurchases. These stocks are familiar to investors in the sense that they are highly visible stocks with large amounts of news circulation, experience frequent tracking by analysts, and produce products often consumed by investors. Extreme Positive Previous Day Return is a dummy variable that takes the value 1 if the stock returns lies in the highest previous day return decile for all stocks listed on the NYSE/AMEX/NASDAQ and 0 otherwise. The extremely positive previous day return dummy captures the impact of the high attention that a stock is likely to receive document in Barber and Odean (2006).

The investor characteristics are: (1) Client segments, (2) Knowledge, (3) Age, and (4) Size to Net worth Ratio. The discount brokerage house classifies its clients into three categories: (A) Affluent Traders: Households with more that $100,000 in equities at any point of time, (B) Active Traders: Households that make more than 48 trades in any year, and (C) General Traders:

Households that not classified as affluent or active traders. Client segments are dummy variables that are coded treating the Active traders as the reference category. The data on their knowledge of investments and net worth is self reported by the investors at the time of opening their accounts. The age dummy is obtained from a demographic database. Knowledge is a dummy variable that takes the value 1 if the investor considers himself/herself to have an extensive or good knowledge of investments and 0 if the investor considers himself/herself to have limited or no knowledge. Size to Net worth Ratio (SNR) is the ratio of the average monthly investment in common stock divided by the self reported net worth of the client. The SNR ratio captures the impact of investors holding a significant portion of their wealth in common stocks invested through the brokerage house. In addition, 70 dummy variables are included as dependent control variables for any time specific clustering of repurchases during the 71 month period January 1991 and November 1996. Household specific preferences for repurchases are largely captured by the client segment classification described above.[11]

The results of logistic regressions are presented in Table 2.9. In Model 1, the dependent variables consist only of stock characteristics. I find that individuals are twice as likely to conduct repurchases in stocks that earn a positive market adjusted net return, which is no surprise given the findings in Section 5.2 relating to investor learning based on performance in prior roundtrip trades. The estimated odds ratios for the technology sector and large stocks show that there is an additional impact of familiarity based on characteristics other than having held a stock before. An odds ratio of 1.22 on highest previous day returns provides support for the attention bias in stocks observed by Barber and Odean (2006). In Model 2, the dependent variables consist only of investor characteristics. The odds ratios indicate that affluent traders are less (i.e. .83

[11] Due to the large number of households, the data becomes intractable for a fixed effect approach aimed at capturing household specific preferences for repurchasing stocks. However, the client segment categorizations, through a broad classification of households, partly capture household specific preferences.

times) likely than general traders to repurchase stocks, while active traders are more than twice (i.e. 2.28 times) as likely to repurchase stocks. Affluent traders are likely to have access to more resources to help them with their investments, which can explain their lower propensity to repurchase in comparison to other traders. Individuals who consider themselves highly knowledgeable seem to repurchase less often in comparison to their less knowledgeable counterparts. However the impact of knowledge is statistically insignificant when stock and investor characteristics are combined in Model 3. Also, knowledge being a self reported subjective variable, has not been used for making far reaching conclusions. Older people seem to conduct fewer repurchases. Individuals tend to gamble with small portions of their wealth. Interestingly, individuals with a higher portion of their wealth invested in common stocks (SNR) seem more disposed to repurchasing. This result does not lend support to the argument that repurchases are largely a gambling activity. In terms of magnitude, among all investor characteristics only the client segment classifications seem to have a big impact on the propensity to repurchase stocks. Model 3 combines all stock and investor characteristics, yielding results similar to previous models that consider investor and stock characteristics separately.

In summary, individuals prefer to repurchase stocks that earn positive market adjusted net returns in their previous round trip trade, are in the technology sector, have a large market capitalization and receive attention based on extremely positive previous day returns. Also, individuals who trade actively exhibit the greatest propensity to repurchase stocks.

## 7 Repurchase: Robustness Check

### 7.1 Are Repurchases mere gambles made in a few select accounts?

It could be argued that the results relating to repurchases might not be relevant as investors are probably just taking a gamble on a few stocks. Shefrin and Statman (2000) state that individuals form layered portfolios where risky gambles are made in an extreme upside potential layer to get a shot at riches. Clearly their retirement portfolio can not be included in this layer. A rational expectation is that if repurchases were mere gambles then we should expect individuals to have a greater likelihood of repurchasing in their non-retirement accounts in comparison to their retirement accounts (IRAs and Keoghs). In percentage terms the composition of repurchasing households that hold only retirement accounts, only non-retirement accounts and both retirement and non-retirement accounts stands at 8%, 42% and 50% respectively. In this section I analyze if repurchases are more likely to occur in non-retirement accounts of households that hold both Retirement and Non-Retirement accounts at the discount brokerage house. On average, 41.26% of repurchases occur in retirement accounts that account for 42.47% of the average monthly common stock positions. For all households on average it seems that repurchases are just as likely to occur in an individual's retirement accounts versus his non-retirement accounts. Table 2.10, Panel A presents the percentage of repurchases and average monthly holding in Retirement and Non-Retirement Accounts for the average household in various familiarity quintiles. Households in the lowest familiarity quintile conduct 41.87% of their repurchases and hold 41.61% of their average common stock holding in their Retirement accounts. These figures are nearly the same for all other familiarity quintiles. The highest familiarity quintile households conduct 38.27% of their repurchases and hold 40.16% of their average common stocks holding in their retirement accounts. Thus there is no evidence that repurchases could be interpreted as gambles irrespective of the degree of familiarity bias among households and seem like a pervasive trading activity.

### 7.2 Do the accounts in which investors conduct repurchases form a significant part of an investor's total wealth and overall common stock trading activity?

Individuals might hold accounts with multiple brokerage houses and thus it can be argued that the trading activity documented in this study might not be of great significance. Goetzmann and Kumar (2005) present evidence from the Survey of Consumer Finance (SCF) data that counters the above argument. According to the 1992 SCF, the median U.S. household held only one brokerage account (mean=1.57) in 1992 (approximately 62% of the households had only one brokerage account) and the 1995 SCF indicates that the median number had increased to two (mean=2.62) in 1995. This indicates that even though some investors might have other brokerage accounts, the majority of them are likely to have only one brokerage account at least during the first half of my sample period.

An alternative to check for the significance of investments in common stock would be to view the investment in common stocks through the brokerage house relative to the total wealth of the client. This study documents the Size to Net worth Ratio (SNR) for all repurchasing households, wherein size refers to the average monthly position in common stocks and Net Worth is self reported by the clients. The data for SNR is available for only 42% of repurchasing households and 41% of all households (i.e., including non-repurchasing households), as not all clients reported their Net Worth at the time of opening their accounts. There does not seem to be any systematic self reporting bias in the reported net worth by individuals at the time of opening their accounts at the brokerage house. Panel B of Table 2.10 presents the percentage breakup of the number of households within each familiarity quintile across SNR breakpoint. The SNR breakpoints are determined from SNR quintiles measured for the entire sample of households (including those that do not make repurchases). The percentage of households within each familiarity quintile having an SNR in excess of 0.48 ranges from the 73% in the lowest familiarity quintile to 59% in the highest familiarity quintile. An SNR of 0.48 is quite high as it implies that 48% of a client's net worth is invested in common stocks at the discount brokerage house.[12] Although these results suffer from data limitations, they still provide some evidence that most households invest a highly significant portion of their total wealth in common stocks through the discount brokerage house.

## 8 Buy and Hold Returns Vs Realized Returns

In this section I explore if the repurchasing strategy outperforms the buy and hold strategy. Investors possessing superior market timing ability might have a greater propensity to repurchase stocks, warranting a comparison of buy and hold returns and realized gross returns. Furthermore, for the repurchasing strategy to be profitability, the gross returns should sufficiently cover commission costs. Commission costs may be a concern for most investors, but these costs become more relevant from the perspective of investors who tend to repurchase at a higher frequency.

A comparison of Buy and Hold Returns, Realized Gross Returns and Realized Net Returns presented in Table 2.11 sheds light on the performance of repurchases vis-à-vis buy and

[12] A SNR greater than 1 implies that that a household might have significant liabilities in relation to its total assets.

hold returns and also highlights whether it is the commission cost or foregone returns in the intermediate period that contribute to any observed underperformance. The methodology involves calculating the returns for every stock that was repurchased by a household over the period beginning when the stock was first sold by a household until the last time it was repurchased. Only those repurchases that take place within 1 calendar year (i.e., 250 trading days) from the last sale of a stock by a household have been considered in these calculations. The returns for a household are calculated based on an average repurchase value weighting of the returns for all stocks repurchased by the household. Lastly, the returns for all households were averaged to determine the returns for each familiarity quintile.

The difference between buy and hold returns and gross returns appears statistically insignificant for the lowest 3 familiarity quintiles. For the fourth and the highest familiarity quintiles the gross returns exceed the buy and hold returns by 3.31 percent and 5.28 percent. It seems that households repurchasing at higher frequencies might be doing so because of their superior market timing ability. With commission costs, the performance for the repurchase strategy versus the buy and hold strategy are expectedly lower. On a net return basis, the buy and hold returns marginally outperform a repurchasing strategy for all familiarity quintiles. However, these differences are statistically insignificant. The impact of commission costs is significantly more for the highest two familiarity quintiles, a result attributable to greater number of roundtrips conducted in the same stock in comparison to other familiarity quintiles. Households in the top familiarity quintile, on average, make approximately two roundtrip repurchase trades in every stock they repurchase, whereas households in the lowest familiarity quintile make only a single roundtrip repurchase trade (refer to Table 2.7, Panel C). Grossman and Stiglitz (1980) argue that people will trade as long the benefit from trading exceeds the cost of these trades.

Contrary to their argument, I find that people repurchase stocks are trading even though the returns on their strategy fail to exceed commission costs. Thus, I conclude that the repurchasing strategy is not superior to a buy and hold strategy, irrespective of a households propensity to repurchase.

## 9 Sub-optimality of Repurchases

### 9.1 Impact of Commission costs on performance of repurchase roundtrip trades

The earlier discussion in Section 4 concludes that the lack of positive abnormal returns associated with gross returns is indicative of repurchases being attributable to familiarity with stocks. This section studies Net Returns measured by adjusting gross returns on repurchase roundtrip trades for commission costs. The results in Table 2.12 show that commission costs further attenuate the performance of repurchase roundtrip trades measured by gross returns. Irrespective of the familiarity quintile the net abnormal returns are significantly negative. In gross return terms, the highest two familiarity quintiles seemed to earn positive market adjusted returns and Carhart 4-factor abnormal returns were insignificantly different from zero. However inclusion of commission costs shows that they earn reliably negative abnormal return based on any benchmark or factor model. The Carhart 4-factor abnormal returns for the highest familiarity quintile in terms of gross returns and net returns are -0.26% (p-value=0.59) and -2.65% (p-value=0.00) respectively. Commission costs further lower the performance of roundtrip repurchase trades by at least 2.5% for any familiarity quintile. Investors may not devote much attention on the impact of commission costs (Barber and Odean (2000)). The evidence shows that repurchase trades, on average, perform poorly and commission costs add to their underperformance. Thus, commission costs warrant serious consideration for investors who make repurchases.

## 9.2 Under-Diversification associated with preference for the familiar

Individuals are largely under-diversified. Blume, Crockett and Friend (1974) find that the average number of stocks in portfolios reported in 1967 by the Federal Reserve Board stood at 3.41. This number has hardly changed with the average number of stocks in a portfolio reported by Goetzmann and Kumar (2005) at 4 during the period 1991-1996. While investing in common stocks, the average household faces information costs and wealth constraints that inhibit adequate diversification. Narrowly focusing on a few stocks they currently hold or held in the past results in further under-diversification. If an individual holds an existing position in a stock, then diversification benefits can be obtained if he or she considers selecting other securities. Irrespective of the degree of familiarity, I find that prior to the day of repurchase, households hold a position in more than a quarter of their repurchased stocks.[13] On the other hand, if an investor is taking a position in a stock that they held in the past but do not own in the present then he/she should consider splitting his/her investment among other securities. In either case the returns after accounting for commission costs do not justify repurchasing the same stock. This section explores whether the degree of familiarity, proxied by the ratio of repurchases to total purchases, has any impact on the observed under-diversification across households.

[13] The average percentage of repurchases with an existing position in the repurchased stocks for the five familiarity quintiles from the lowest to the highest are 28.9, 26.49, 26.04, 25.43 & 28.90, respectively.

### 9.2.1 Diversification Measures

In order to explore if familiarity bias is associated with the formation of sub-optimal portfolios, this study uses five different measures of diversification.[14] The first measure used is the number of stocks (STK). Although the STK measure does not account for correlation among securities, it is the simplest measure and has been used in several other studies (Blume, Crockett and Friend (1974) and Vissing-Jorgensen (1999)). A higher STK is likely to indicate greater diversification.

The second measure considered is the sum of weighted squares (WSQ) used in Blume and Friend (1975). It is calculated as:

$$\text{WSQ} = \sum_{i=1}^{N} \left(w_i - w_{im}\right)^2 \approx \sum_{i=1}^{N} w_i^2$$

where $w_i$ is the proportion invested by a household in security i, $w_{im}$ is the proportion of security i in the market portfolio and N is the number of securities. WSQ attempts to determine how closely a household replicates the market portfolio and is measured by summing the squared deviation of the proportions invested in each security from the proportion in the market portfolio. Since the proportion in the market portfolio is small, this measure can be approximated as the sum of the square of the proportion invested in each security. A lower WSQ measure indicates greater diversification.

The third measure of diversification is the number of industries across which households diversify their holdings (IND). All stocks are classified across 10 broad based industry classifications.[15] Although this measure does not capture the exact correlation among stocks, it is a better indicator of diversification than the STK measure. A higher IND measure indicates greater diversification.

[14] Except for the industry measures, these measure can be found in Goetzmann and Kumar (2005)
[15] The industry definitions have been obtained from Kenneth R. French's online data library.

The fourth measure considered is the industry weight square (IWSQ) measure and is calculated as:

$$\text{IWSQ} = \sum_{ind=1}^{N} w_{ind}^2$$

where $w_{ind}$ is the proportion invested in each industry and N is the number of industries. IWSQ measures the concentration of investments across households and is similar to the Herfindahl-Hirschman Index (HHI), a commonly accepted measure of concentration of firms within an industry. A lower IWSQ shows lower concentration of holdings within few industries, indicating a better diversification strategy.

The last measure is the normalized version of the portfolio variance (NV) used in Goetzmann and Kumar (2005).[16] The expected portfolio variance of an equal weighted portfolio of N stocks is given by:

$$\sigma_p^2 = \frac{1}{N}\overline{\sigma}^2 + \left(\frac{N-1}{N}\right)\overline{\text{cov}}$$

where $\overline{\sigma}^2$ is the average variance of all stocks in the portfolio and $\overline{\text{cov}}$ is the average covariance among all stocks in the portfolio. The covariance matrix for each portfolio at the end of each month is calculated using the previous five years of monthly returns.[17] The normalized variance (NV) is obtained by dividing the portfolio variance by the average variance of stocks in the portfolio:

$$NV_{ewp} = \frac{\sigma_p^2}{\overline{\sigma}^2} = \frac{1}{N} + \left(\frac{N-1}{N}\right)\left(\frac{\overline{\text{cov}}}{\overline{\sigma}^2}\right)$$

[16] This measure was proposed by Goetzmann, Li and Rouwenhorst (2005).
[17] Only those stocks with more than 24 months of data within the last 5 years are included for estimation of the covariance matrix.

This normalized variance provides a meaningful average diversification estimate when a set of portfolios of different sizes are examined. The above expression for NV indicates that there are two ways of decreasing the portfolio variance. Firstly, it can be reduced by increasing the number of stocks in a portfolio as indicated by the first term (1/N) in the equation. Secondly, it can be reduced by selecting stocks with lower covariance (or correlations). Simply increasing stocks in a portfolio without considering the correlations between stocks is referred to as "naïve" diversification. On the other hand, picking a portfolio with a lower average correlation among stocks reflects an understanding of the advantage of portfolio theory and is referred to as "skillful" diversification.

### 9.2.2 Impact of familiarity on diversification

This section studies the diversification measures observed across various familiarity quintiles. Table 2.13, Panel A shows that households with a greater familiarity bias tend to hold a fewer number of stocks. The lowest familiarity quintile holds an average 8.14 stocks while the highest familiarity quintiles hold an average of 2.83 stocks. A monotonic increase in the WSQ measure shows that households with greater familiarity bias tend to be more under-diversified in comparison to the market portfolio. In terms of number of industries, the lowest familiarity quintile households tend to hold stocks in two more industries than the highest familiarity quintile households. The monotonic decrease in the IND measure across familiarity quintiles indicates that households with a lesser degree of familiarity bias display evidence of skillful diversification. The IWSQ measure shows that households with greater familiarity bias tend to concentrate their holding in a few industries.

Panel B presents the results for the normalized variance (NV). The same trend in diversification is observed here as well with the highest familiarity quintiles showing a significantly larger normalized measure in comparison to the lowest familiarity quintile. The NV naïve (NVN) measure refers to the first term (1/N) in the equation presented in the previous section. The results show that a significant part of diversification among households with a lesser degree of familiarity bias can be attributed to simply investing in more stocks irrespective of the correlation structure among them. The last measure presented is the average correlation measure for portfolios that hold at least 2 stocks. This measure provides an insight into the diversification benefits that accrue from "skillful" diversification. The correlation among stocks gradually declines with higher familiarity quintiles and range from 0.2542 for the lowest quintile to 0.2778 for the highest quintile. The numbers, though statistically significant, might not appear to be economically significant if the average number of stock holdings across familiarity quintiles is ignored. Maintaining an average correlation of 0.2542 among 8 stocks certainly indicates greater skill than maintaining a slightly higher correlation of 0.2778 among 3 stocks. Combining this result with the observed diversification across industries, there is evidence of less familiar households being more "skillfully" diversified.

This evidence leads to the question of why the average household with lesser familiarity bias should posses more skill. The answer lies in the composition of households within each familiarity quintile. As mentioned before, the discount brokerage house at which the investors maintain accounts categorizes households into three categories: (A) Affluent Traders, (B) Active Traders, and (C) General Traders. Table 2.14, Panel A shows the percentage composition of various traders across each familiarity quintiles. The proportion of active and affluent traders as a percentage of all investors tends to decrease with higher familiarity quintiles. In other words, households with greater levels of familiarity bias consist of a greater proportion of general traders. Panels B & C show the average number of stocks and number of sectors held by households within each familiarity quintile and within each investor category. It seems like the General Traders do a poor job of diversifying across multiple stocks and sectors. Active traders are likely to be more diversified because of their active interest in stocks and affluent traders might possess access to more resources (like financial planners and stock brokers) enabling them to be more diversified. The evidence indicates that the under diversification observed amongst higher familiarity quintiles can be ascribed to the larger presence of general traders, who are probably the least sophisticated of all three categories. The non-repurchasing households tend to have a large composition of general traders, accounting for the low level of diversification observed.

Overall, the evidence indicates that households with more familiarity bias tend to hold more under-diversified portfolio. A look at some measures like STK and IND shows that most households tend to be largely under-diversified. Statman (1987) finds that a well-diversified portfolio of randomly chosen stocks must include at least 30 stocks. In comparison, an average of 8.14 stocks for the lowest familiarity quintile falls well short, indicating that households largely remain under diversified. Also, even though investing in 4 sectors seems a lot better than investing in 2, it still shows that 6 other broad sector categories are largely ignored. Goetzmann and Kumar (2005) find that due to the perceived familiarity with local stocks, investors with stronger local bias (another aspect of familiarity bias) hold relatively less diversified portfolio. This study provides another dimension to the observed under-diversification associated with preferring familiar stocks.

### 9.2.3 Are repurchasing households adequately diversified with their mutual fund holding?

It could be argued that the observed under-diversification for repurchasing households could be unimportant as these households might be adequately diversified through their mutual fund investments. The evidence presented in Section 7.2 argues that the investment in common stocks at the brokerage house forms a significant portion of an average repurchasing household's wealth. Figure 2.2 presents the breakup of investments made by repurchasing households among common stocks, mutual funds and other investments (eg: Treasury bills, Options etc). The lowest familiarity household has 72.78% invested in common stocks and 14.25% invested in mutual funds. The ratio of investments in common stocks to mutual funds increases monotonically as we move to higher familiarity quintiles, with the highest familiarity investing 80.61% in common stocks and 10.04% in mutual funds. Thus, there is little support for the argument that investors with a higher familiarity bias might not be more under-diversified as they might have achieved their diversification objectives through mutual funds.

## 10 Summary and Conclusion

Analysis of the trading records of investors at a large brokerage household shows that 40% of the households, representing 80% of the value of purchase trades made by all households, repurchase a stock previously sold. The gross returns for these repurchases, measured over roundtrip trades across various groups, shows that repurchasing households are unable to outperform the market or industry benchmarks. Also, the abnormal return measures are reliably insignificant for all repurchasing household categories. The evidence indicates that repurchasing activity amongst households is not driven by possession of superior or private information in a semi-strong efficient market. Investors fear the unknown and therefore prefer assets that are familiar to them, referred to as the familiarity bias.

The familiarity bias does not operate in isolation. An investor's experience with a stock in the roundtrip trade prior to repurchase significantly impacts his or her decision to repurchase. Irrespective of the degree of familiarity bias, investors are more willing to bet on stocks that performed comparatively well in prior roundtrip trades. Investors seem to view past roundtrip returns in a stock as representative of the future. More specifically, the average investor makes an additional roundtrip repurchase transaction on a stock only if he or she is able to outperform the market after adjusting for commission costs in the prior roundtrip trade. Benartzi (2001) finds similar evidence of extrapolation of past returns by individual investors in retirement savings allocation as well. In addition, this study finds that repurchasing households tend to hold on to their losing stocks (i.e. non-repurchased stocks) longer, in comparison to their winning stocks (i.e. repurchased stocks). One of the central findings of this study is that familiarity leads to an individual being pre-disposed to repurchasing a stock and the decision to repurchase is positively related to performance of a stock in its prior roundtrip trade.

The repurchasing activity of households seems to be concentrated in a few highly visible firms with large amounts of news circulation, experience frequent tracking by analysts, and produce products often consumed by investors. Thus it is no surprise to see investors tilt their repurchases to these stocks as they probably feel more comfortable with these familiar stocks. In the sample of households, as much as 36% of the repurchases are made in the technology sector, attributable to the euphoria surrounding technology stocks in the 1990s. Also, individuals exhibit a preference for repurchasing stocks of companies that experience extreme positive returns on the previous day, a confirmation of the attention bias documented in Barber and Odean (2006).

Individuals who trade very actively are more likely to repurchase more frequently than other individuals. I find no evidence for the argument that repurchases are a gamble investors take in accounts that account for a small portion of their total wealth.

Individuals following a repurchasing strategy fail to outperform buy and hold returns. Another important conclusion of this study is that repurchasing stocks is a sub-optimal trading activity. First, like Barber and Odean (2000), I find that commission costs significantly lower the performance of repurchase. In particular, the highest familiarity households earn insignificant alphas on a gross basis but including commission costs make these alphas significantly negative. Lastly, households with a greater familiarity bias tend to be more under diversified. Households with a lower portion of repurchases in their total purchases (i.e., a lower degree of familiarity bias) tend to hold more stocks with lower average correlations and invest across more sectors in the economy. In other words, households with lower familiarity bias display superior diversification skills. I find that active and affluent traders, representing a lower proportion of households in the high familiarity bias group, tend to be better diversified than general traders. Investors with a greater degree of familiarity bias invest a larger portion of their wealth in common stocks and seem to reduce their investments in mutual funds. Thus there is little evidence of investors achieving greater diversification benefits through mutual funds.

The results of this study show that individual investors should do better for themselves by focusing on stocks beyond those held in their portfolio and considering the adverse impact of trading costs and foregone returns while frequently repurchasing stocks.

# CHAPTER 3

# INDIVIDUALS AND THEIR LOCAL UTILITY STOCKS: PREFERENCE FOR THE FAMILIAR

---


## Abstract

*I find individuals are nearly four to five times more likely to purchase stocks of their local direct utility as opposed to utility companies operating outside their state of residence. My tests reveal that individuals do not possess superior or private information about their local utilities, nor are they using their local utility stocks as a hedge for possible increase in their utility expenditure. Indeed, individual preference for their local utility stocks seems to be driven by preference for familiar assets, referred to as familiarity bias. In addition, I find that this pervasive behavior can not be attributed to individuals who are less affluent or less sophisticated than their counterparts.*


## 1. Introduction

Investors often prefer assets they are familiar with, referred to as familiarity bias. They have a dislike or fear of things that they are unfamiliar with. This study provides evidence of familiarity bias in the geographical context, focusing on an individual's decision to purchase utility stocks. The utilities sector provides a unique setting to analyze the impact of familiarity on investment decisions made by individuals. In the United States, most sectors of the economy have a few large corporations in which the majority of the trading activity occurs and these corporations have operations or distribution networks spread throughout the length and breadth of the country. It seems unlikely that for these sectors, familiarity associated with company geographical location of the companies would impact an individual's decision to purchase. In other words, the big familiar stocks are national (even global) in operations. However, the utilities sector is unlike most sectors of the economy. The operations of utility firms tend to be more local or regional in nature. Thus, investors tend to be familiar with the utility firms in their region, and not so familiar with distant utility firms. For example, most New York residents would be unaware of Avista Corp., a utility provider in the Pacific northwest of the United States. Moreover, I study only those utility firms that directly interact with individuals through their distribution networks, referred to as direct utility providers.[18] In addition, information asymmetries are likely to be higher for direct utility stocks, since less than 20 percent of these stocks in my sample are listed on the S&P 500, which in turn results in lower analyst and media coverage.

In January of 1942, John Maynard Keynes purchased a large position in Elder Dempster for The Provincial Insurance Company. F. C. Scott of Provincial questioned why such a large position was purchased, Keynes replied "…that I preferred one investment about which I had sufficient information to form a judgment to ten securities about which I know little or nothing." (Moggridge, 1983) Keynes' response illustrates the cognitive resource limitations of learning about the many firms needed to be diversified and possibly hints at a common heuristic adaptation, familiarity. Familiarity bias associated with geographical location of businesses has been documented by other studies as well. For example, Deogun (1997) discusses the stupendous growth of the Coca Cola Company from a stock valuation of $4 billion to $145 billion over a span of 16 years and states that, "Though the company's reach is global, the rewards are magnified in Atlanta because so much of the stock ownership is local. At least $23 billion of Coke stock, or 16%, is held in Georgia, most of it in metropolitan Atlanta, and for many

[18] I exclude firms that are purely wholesalers of utilities since they never directly interact with their customers. For example, PG&E appears on customer billings for California residents, but these individuals would generally be unaware of wholesaler that PG&E procures gas from.

shareholders, selling is anathema." Huberman (2001) considers the geographical distribution of the shareholders of the seven U.S. Regional Bell Operating Companies (RBOCs) at the end of 1996 and finds that in most states, more money is invested, per investor, in the local RBOC than in any other RBOC. He finds no support for the argument that a customer of an RBOC may over-invest in its stock as a hedge against an unexpected increase in the price of its services. This familiarity bias contradicts traditional portfolio theory, which implies that investors should diversify and invest less in the RBOC serving him than in those other parts of the country since the fortunes of the RBOCs vary with the economic tides in their home areas. Grinblatt and Keloharju (2001) in their study of the Finnish stock market find that investors simultaneously exhibit a preference for nearby firms and for same-language and same-culture firms.

During the initial exploration of the trading data of individual investors at a large discount brokerage house between January 1991 and November 1996, I found that the concentration of purchases among the top three firms within each of the 49 industrial sectors was the lowest for the utility sector. For example, the top three firms, ranked according to the portion of purchases at the brokerage house for the non-alcoholic beverage stocks (Coca Cola, Pepsico & Wrigley William Jr. Co) was 92.37 percent, in comparison to only 12.77 percent for utility stocks (PG&E, Texas utilities, Edison International). Clearly, individuals are not focussing on a few large stocks within the utility sectors, unlike they do in some other industry sectors. It turns out that these investors concentrate their purchases in local utility stocks. In the first part of this study, I provide evidence of preference for local utility stocks. My results show that individuals residing across all four major geographical divisions in the United States are likely to allocate significantly more than their expected allocation (i.e. allocation in the absence of any location bias) to their local utility stocks. Modeling the decision of utility investing households to buy each of the 170 utility stocks in my data, I find that after controlling for various other factors, households are 4 to 5 times more likely to buy their local utility stocks than their non-local utility stocks.

Other than individuals being familiar with their local firms, there exist alternate explanations for individuals preferring their local utility stocks: (1) individuals could possess private or superior information about their local utility stock, or (2) individuals could be using their long positions in their local utility stocks to hedge themselves against an increase in utility prices. Coval and Moskowitz (1999) found that U.S. investment managers, in a setting of a single currency and relatively little geographical variation in regulation, taxation, political risk, language, and culture, prefer to hold companies headquartered close to them. Their results suggest an information based explanation for local equity preference since the firms preferred by the investment managers tend to be small and highly levered, and they tend to produce goods not traded internationally (i.e. firms with greater information asymmetry). Massa and Simonov (2006) found that Swedish investor stock holdings were driven by geographical and professional familiarity and that this familiarity is not a behavioral bias, but is information driven. Ivkovic and Weisbenner (2005), using the same brokerage house data as ours, find that individual investment in local firms is information based for the entire cross-section of industries in the US. This is in stark contrast to my study, which looks at only direct utility firms, where information asymmetries are likely to be higher, thus providing greater opportunities to benefit from private information. Yet, I reject the information-based explanation.

If the local bias in utility stocks is information based, we would expect the following: (1) Local utility investments should outperform non-local investments, and (2) Local utility investments should earn positive abnormal returns. I find that neither of these conditions hold true. Furthermore, to test any information advantages within local utility stocks, I made comparisons of Non-S&P 500 Vs S&P 500 stocks and Regional (i.e. firms operating in one state) Vs Non-Regional stocks (i.e. firms operating in multiple states). Non-S&P 500 stocks receive less attention from analysts and the press in general relative to S&P 500 firms. If individuals can access some superior or private information about these firms due to their geographical proximity, investments in these stocks are likely to yield positive abnormal returns and outperform their S&P 500 counterparts. An information based story would have similar implications for the Regional Vs Non-Regional stocks comparison. And yet, I find that the abnormal return performance of roundtrip trades did not reflect the presence of superior information.

Massa and Simonov (2006) find that Swedish investors do not hedge, but instead, invest in stocks closely related to their non-financial income. To check if investors were holding long positions in utility stocks to hedge themselves from an increase in utility prices, I tested for the presence of a positive correlation between returns on utility stocks and the percentage changes in utility prices. Interestingly, I found that the correlations were negative or insignificantly different from zero, providing no support to the hedging argument. It should also be noted that utility firms tend to be heavily regulated, preventing them from increasing their product prices easily. Prior studies have provided explanations, other than the familiarity bias, to explain the home bias puzzle. Serrat (1997) has shown that in an international exchange economy with two agents, the home bias puzzle can be attributed to the non-tradability of some goods affecting the marginal utility of tradable goods. Stulz (1981) provides evidence on how restrictions on international capital flows could lead to the preference for domestic assets. However, the same arguments cannot be used to explain the impact of familiarity in domestic investments.

Overall, I fail to find support for the information and hedging explanations, leading us to conclude that local utility investments are driven by geographical familiarity with firms. Leng and Seasholes (2005) and Chen et al. (2007) show that investor sophistication and wealth has some impact on behavioral biases. One could argue that individuals who buy their local utility stocks are simply the less affluent or unsophisticated investors, in which case my findings provide us with little understanding of individual investor behavior as I could be depicting the behavior of "noise" traders. Contrary to expectation, I find that the average local utility stock investor is as, or marginally more, affluent or sophisticated than the average individual investing solely in non-local utility stocks, suggesting that sophistication and affluence might have little to do with preference for familiar assets. Investor sophistication, reflected in portfolio diversification, shows that the small portion of investors holding more than ten stocks in the portfolio generate positive abnormal returns. Thus, highly sophisticated investors are more likely to make sound investment decisions, which are likely driven by some superior information about their local utility as opposed to pure familiarity.

## 2. Data

The data covers the holdings and trading activity of households at a large discount brokerage house during the period January 1991 to November 1996, details of which are provided in Odean (1999) and Table 2.1. In these accounts, there were 38,872 ($339 million) purchases made in the utilities sector by 16,578 households in the data set. These households represented 26.34% of households that traded in stocks during that period. The purchase trades in the utility sector represented 3.78% (2.93% in value terms) of all the purchases made at the discount brokerage house. There were 35,689 transactions made by 11,343 households involving almost all (except

3) direct utility stocks. The Infobase database provides the zip codes for residence for about 60% of the households that traded in utility stocks. My analysis is restricted to only households for which zip codes are available. The demographic information containing the exact zip code is available for 7,149 households and this allows us to analyze 21,794 utility stock purchases. Out of these, 6,183 purchases were made in stocks that provided utility services in the investor's state of residence.

There were 239 utility firms listed on the NYSE/AMEX/NASDAQ during the 1991-1996 period that have been classified into two categories: (1) Direct Utility firms, and (2) Indirect Utility firms. The former refers to 177 firms involved in direct distribution of utilities, while the latter refers to 62 firms that supplied utilities to other marketing and distribution firms. This study focuses solely on purchases made in direct utility firms for two reasons. First, direct utility firms interact with individuals through their distribution network, while the indirect utility firms supply utilities to other firms for further distribution to meet various end user needs. Generally, the end user only knows about the direct utility firms and is unaware of the original suppliers. Second, unlike the indirect utility firms, most of the direct utility firms are regional corporations with operations in only one or two states. Due to this regional focus, individuals are likely to be more familiar with their regional direct utility providers vis-à-vis other direct utility firms. Approximately 90 percent of direct utility providers consisted of electricity and gas providers. I now use the term "utility" as synonymous with "direct utility".

To capture the extent of a utility provider's interaction with households, I use the firm's areas of operation, which is unavailable in any financial database. The Compustat database provides the location of the registered office and headquarters of a firm. Information on registered offices is not suitable for my analysis as quite a few firms are incorporated in

Delaware primarily for taxation purposes. Since a significant portion (i.e. 42 percent) of direct utility firms operate in multiple states, headquarter locations by themselves will not accurately depict the extent of a firm's interaction with households. The data on the areas of operation and type of business operations (i.e. direct or indirect utility provider) for these firms has been hand collected from Form 10K and Annual Report to Shareholders fillings with the SEC. I refer to either the last statement filed prior to the last time a stock is traded or the last statement filed for a fiscal year prior to 1997, whichever one is earliest. The Form 10K and Annual Report to Shareholder follow a fixed format laid out by the Securities Exchange Commission (SEC). However the description of business varies, with most companies providing information that does not go beyond their specific states of operation. Given the constraints, I confine my data to the states of operation for all direct utility firms.

I find that the households at the discount brokerage made 28 percent of their direct utility stock purchases in utility companies operating within their state of residence. The U.S has 50 states, so for investors to concentrate such a large portion of their purchases within their local utility stocks is surprising. If there was pure randomness in the geographical context, we should expect stock purchases of utility firms having operations in the investor's state of residence to approximate 2 percent.

Figure 3.1(a) presents a distribution of purchase transactions based upon the distance[19] between the area where an individual resides and the closest area of operation for the utility stock purchased by the individual. To get a better sense of distances involved, note the distances

[19] For every transaction, I measure the great circle distance (in miles) between the centroid of the household's zip code and the centroid of the nearest state of operation for the direct utility stock purchased. The data on latitude and location of each zip code and states centroid are available on the SAS website, a statistical software provider. The great circle distance is often used in geographical and aviation studies, and is calculated as:

$$\textit{Great Circle Distance (miles)} = 3949.99 \times \{acos\ [(sin(Y_2)\ sin(Y_1)) + (cos(Y_2)cos(Y_1)cos\ (X_2\text{-}X_1))]\}$$

where, $Y_1, Y_2, X_1$ and $X_2$ refer to radian measure of the latitude for location 1, latitude for location 2, longitude for location 1 and longitude for location 2, respectively.

between four major cities in the US: (1) Distance between Seattle, WA on the west coast to New York City, NY on the east coast is 2,401 miles, (2) Distance between Minneapolis, MN in the mid-west and San Antonio, TX in the south is 1,107 miles. I find that 37 percent of purchases in direct utility stocks are made within a 300 mile radius. Nearly half the purchases are made within a 600 mile radius. The distribution based upon the value of purchases is presented in Figure 3.1(b) and provides similar evidence. Utility stocks operating within a 300 mile radius account for 35 percent of purchases. I do not use distance as a measure of local bias any further in my study as I not know the exact county/zip codes for areas of operations for direct utility firms, which induces noise in the distance measure.

## 3. Methodology

### 3.1. Return and Performance Measurement

The household trading records provide the prices at which trades take place along with the commission costs. I use these prices to calculate roundtrip trade returns. A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock. I prefer the use of roundtrip trade returns over returns based on month-end utility portfolio holdings for a variety of reasons. The households in my dataset generally hold a position in only one utility stock and these positions exist only for a few months during the 71 month span, which hinders calculation of factor based abnormal returns using the month-end portfolio approach. Moreover, the roundtrip trade returns provide the exact realized returns on trades.

The return computations are made in three steps. First, the gross returns for a roundtrip transaction j on security i purchased on day 1 and sold on day T is calculated as:

$$R_j^{Gr} = \left[ \prod_{t=1}^{T} (1 + R_{i,t}^{Gr}) \right] - 1,$$

where the daily returns for all days except t=1 and t=T are obtained from CRSP. The gross and net returns on day of purchase ($R_{i,1}$) and day of sale ($R_{i,T}$) are calculated as:

$$R_{i,1}^{Gr} = P_{i,1} / P_{i,b} \quad \text{and} \quad R_{i,T}^{Gr} = P_{i,s} / P_{i,T} \,,$$

where $P_{i,1}$ is the closing price on day of purchase, $P_{i,b}$ is the purchase price, $P_{i,s}$ is the sale price and $P_{i,T}$ is the opening price on day of sale. Second, after obtaining the roundtrip trade returns, I convert each realized return to a monthly return. Third, the gross ($R_h^{Gr}$) monthly returns for roundtrip trades made by a household h during the period January, 1991 to November 1996 is calculated as:

$$R_h^{Gr} = \sum_{j=1}^{n_j} w_j R_j^{Gr} \,,$$

where $R^{Gr}_j$ refers to the gross monthly return on round trip transaction j, and $w_j$ refers to the dollar value of transaction j scaled by the dollar value of all transactions made by household h. In the last step gross returns ($R_k^{Gr}$) for all households assigned to a particular category are calculated as:

$$R_k^{Gr} = \frac{1}{n_k} \sum_{h=1}^{n_k} R_h^{gr}$$

where $n_k$ is the number of households in a particular category k. The methodology for calculating value weighted returns for each household and the average returns for households is similar to Barber and Odean (2000).

### 3.2. Abnormal return performance

In this study, six measures of abnormal returns for roundtrip trades are considered. The first measure is the market adjusted return, obtained by subtracting the market return from the roundtrip trade return. The CRSP value weighted NYSE/AMEX/NASDAQ index is used to proxy for market returns and the return on this index is measured over the duration of each roundtrip trade.

The second measure is the utility industry adjusted return, obtained by subtracting the utility industry return from the roundtrip trade return. The utility industry return is calculated as the value weighted return of the utility industry excluding the utility stock being analyzed and is measured over the duration of each roundtrip trade. This measure reflects on the performance of a utility stock trade relative to its peers. For calculating the utility industry return, I use the definition under the 49 industry classification, available on Kenneth French's website.

The third measure used is Jensen's alpha (α):

$$\alpha_{j,t} = (R_{j,t} - R_{f,t}) - \hat{\beta}_{j,t}(R_{m,t} - R_{f,t})$$

where $\alpha_{i,t}$ is the market model abnormal return for a roundtrip transaction j that is conducted over t calendar days, $R_{i,t}$ is gross or net return on the roundtrip transaction, $R_{f,t}$ is the one month Treasury bill rate scaled to time t. $R_{m,t}$ is the market return on the CRSP value weighted NYSE/AMEX/NASDAQ index measured over time duration of t calendar days, and $\beta_{i,t}$ is the beta in the CAPM. To adjust for any impact of non-synchronous trading the beta is estimated using the Scholes and Williams (1977) procedure over a time window of 250 trading days prior to each transaction.

The fourth measure uses the Fama and French (1993) 3-factor abnormal return and is calculated as:

$$\alpha_{j,t} = (R_{j,t} - R_{f,t}) - \hat{\beta}_1(R_{m,t} - R_{f,t}) - \hat{\beta}_2(SMB) - \hat{\beta}_3(HML)$$

where SMB is a size factor, HML is a value factor and the description for remaining terms is the same as in the CAPM described above.

The fifth measure used is the obtained from the Carhart (1997) 4-factor model and is calculated as:

$$\alpha_{j,t} = (R_{j,t} - Rf) - \hat{\beta}_1(R_{m,t} - R_{f,t}) - \hat{\beta}_2(SMB) - \hat{\beta}_3(HML) - \hat{\beta}_4(WML)$$

where the WML is a momentum factor and the description for remaining terms is the same as in the Fama and French 3-factor model.

The last measure uses a five factor model that adds an industry factor to the Carhart (1997) 4-factor model and is calculated as:

$$\alpha_{j,t} = (R_{j,t} - Rf) - \hat{\beta}_1(R_{m,t} - R_{f,t}) - \hat{\beta}_2(SMB) - \hat{\beta}_3(HML) - \hat{\beta}_4(WML) - \hat{\beta}_5(IND)$$

where IND is the utility industry return calculated as a value weighted return of all utility firms, except the utility stock being analyzed. The IND factor is used to capture any industry effects that might be driving utility stock returns. The coefficients for all the factors in the Fama and French (1993), Carhart (1997) and the five factor model are measured over a time window of 250 trading days prior to the repurchase trade. The data for $R_f$ (1 month Treasury Bill rate), SMB, HML and WML factors is obtained from Kenneth French's online data library.

## 4. Analysis of allocation to local utility stocks at the household level

This section explores whether households overweigh their local (in-state) stocks as a proportion of their entire utility stock portfolio, both in terms of the number and value of stocks. According to the US Census 1990 classification, I divide all households into the following four regions: Midwest, Northeast, West and South. I compute a household's observed local utility allocation in value (or number of stocks) terms by averaging the percentage of the total value (or number of stocks) of its utility portfolio invested in local (in-state) utility stocks over the number of months in which it held utility stocks. The expected allocation for each household in value terms is equal to the market capitalization of all companies operating within the state by the total market capitalization of all utility stocks in the US. Similarly, the expected allocation in number terms for each state is equal to the number of all companies operating within the state divided by the number of utility stocks in the United States. The average observed and expected allocation for each region in obtained by averaging out the observed and expected allocations for households residing in various regions.

I first test out the hypothesis that the observed allocation to local utility stocks exceeds the expected market allocation for households in each region.

$$H_0 : Observed\ Allocation\ (A) = Expected\ Allocation\ (B)$$
$$H_a : Observed\ Allocation\ (A) > Expected\ Allocation\ (B)$$

The results indicate that for the majority of the states in my sample, the observed allocation (both in terms of number and value) to local stocks exceeds the expected allocation. Table 3.1, Panel A shows that the West Region with the maximum number of households in my sample has an average observed and expected allocation in value terms of 20.7 and 4.3 percent, respectively. The excessive allocation of 16.4 percent to local utilities for the West is statistically significant (single sided p-value < 0.01). I find a similar level of statistical significance for other regions both in terms of value and number of stocks (Panel B). The evidence favors the alternate hypothesis, which states that households allocate a greater portion of their portfolio to local utilities in comparison to the expected allocation.

To further test the robustness of the above results I conducted a stronger test to see if the observed allocation to the local stocks is in excess of four times the expected allocation:

$$H_0 : Observed\ Allocation\ (A) = 4 \times [Expected\ Allocation\ (B)]$$
$$H_a : Observed\ Allocation\ (A) > 4 \times [Expected\ Allocation\ (B)]$$

The last column in Table 3.1 shows a highly statistically significant (single sided t-test p-value<0.01) result for the alternate hypothesis in all four regions, irrespective of whether the allocation is measured in terms of value or number of stocks. For example, the paired t-test value for the West Regions in terms of value and number of stocks is 7.18 and 13.64 respectively. The evidence states that the observed portfolio allocation to local utility stocks for households at the discount brokerage house exceeds four times their expected allocation.

In order to test the robustness of my results, I classified the households across nine geographical divisions[20] as per the 1990 US Census classification and found similar results. Only two exceptions were found while conducting the second test for observed allocation being in excess of four times the expected allocation. These were the New England and East South Central Division, which has the fewest number of households in my sample. In addition, I further conducted tests at the state level and found that the observed household allocation across the majority of the states exceeded four times their expected allocation. In the interest of conserving space, I have not presented these additional results for geographical classification at the divisional and state level. Overall I conclude that households, irrespective of the type of geographical classification, allocate a far greater portion of their portfolio to their local (in-state) utility stocks than expected.

[20] These nine divisions were: East North Central, East South Central, Middle Atlantic, Mountain, New England, Pacific, South Atlantic, West North Central and West South Central.

## 5. Impact of location in decision to invest among various utility stocks

This section involves a firm level analysis using logistic regressions to model the impact of location on a household's decision to purchase one of the 170 direct utility stocks.[21] Thus, I conduct 170 individual logistic regressions modeling the decision to purchase each of the utility providers, which were purchased by at least one household during the 71 month period. Each regression has 7,149 observations, representing all the households that made at least one purchase in any direct utility. Only the first purchase made by a household in each stock is considered in each regression to remove possible impacts of clustered observations. The average estimates for all 170 regressions for each regression model are presented in Table 3.2. Each Model presents estimates for all regressions that converged during maximum likelihood estimation in two sub-parts. Part (A) presents all the estimates, while Part (B) presents estimates after replacing the insignificant (α=.05) estimates with a zero.

I use the following logit specification for modeling a household's decision to purchase a stock x:

$$\log\left(\frac{\prod_x}{1-\prod_x}\right) = \alpha + \beta_1 (Local\ Dummy) + \sum_{f=2}^{n} \beta_f F_f$$

where, $\prod_x$ refers to the probability to purchase a stock x, Local (In-state) Dummy acquires a value of 1 if the household resides in the same state as the operational location of the company and 0 otherwise, $F_f$ refers to a specific household characteristic or investment condition and n refers to the number of independent variables in the regression.

[21] There are a 177 utility stocks listed on NYSE/NASDAQ/AMEX, of which 3 were never by any of the households in my sample. In addition, 4 stocks were sold but never purchased by any household during the 71 month period. Thus, I confine my analysis to 170 utility stocks that were purchased at least once by any of the households.

Model 1 includes only Local Dummy as an independent variable. The average of the estimate on the Local Dummy, $\beta_1$ for Model 1(A), representing 154 converged regressions out of 170 regressions, is 1.38 (refer Table 3.2, Panel B). This estimate translates to an odds ratio of 3.98, implying that a household is nearly four times as likely to buy its local utility stock as opposed to a non-local utility stock. For Model 1(B) the results appear stronger with an average estimate of 1.54 (odds ratio of 4.67). These results are supported by the median estimates (refer Table 3.2, Panel C) of 1.62 and 1.5 for Model 1(A) and Model 1(B), respectively.

An average estimate might not be a true representation of all the estimates obtained for each regression model. Therefore, I conduct a one sided binomial test (non-parametric test) to see if the proportion of positive/negative (depending on the sign of the mean coefficient) coefficient estimates ($\beta_1$) for the Local Dummy variable (dummy variable indicating same state of household residence as firm) in each of my regression coefficients is significantly greater than 0.5. The proportion of positive $\beta_1$ estimates in Model 1(A) out of all estimates equals 0.93 (significantly greater than 0.5 at $\alpha$=0.01). In addition, the proportion of positive $\beta_1$ estimates after replacing the insignificant estimates with zero equals 0.72 (significantly greater than 0.5 at $\alpha$=0.01).

Individuals often take the past performance of a stock as reflective of the future, also referred to as the extrapolation bias (Lakonishok, Shleifer and Vishny, 1994; Benartzi, 2001).[22] Model 2 includes the prior one month industry adjusted return to control for individuals picking stocks that have performed well in the past month. A mean coefficient of Model 2(A) of -0.04

[22] Lakonishok, Shleifer and Vishny (1994) provide evidence that value strategies yield higher returns than naïve strategies pursued by other investors because they exploit sub-optimal behavior of investors and not because they are fundamentally riskier. Amongst other explanations, they state that naïve strategies might involve extrapolation of past earning too far into the future, a prediction of the extrapolation hypothesis. Research studying the sub-optimal behavior of employees concentrating a large part of their 401(k) assets in the company stock also reports that this behavior is exacerbated when the company's stock has recently performed well (Benartzi, 2001).

implies that contrary to my expectation, individuals show a tendency to pick stocks underperforming their industry counterparts. However, after replacing the insignificant coefficients with zero in Model 2 (B), I obtain a median coefficient of zero and results of the binomial test in Table 3.2, Panel D (only a significant 33 percent of coefficients are less than zero) show that there is no significant evidence supporting individuals following a specific trading strategy with their utility stocks. The Local Dummy remains significant with a mean and median of 1.53 (odds ratio of 4.6) and 1.46 (odds ratio of 4.3), respectively.

The discount brokerage house classifies households that maintain an average account balance of $100,000 at the brokerage house as Affluent. I incorporate an Affluent Investor Dummy variable in Model 3 to capture any heterogeneity amongst households. The average coefficient of -0.04 for the Affluent Investor Dummy indicates that the wealthy investors are less likely to purchase their local utility stocks. However, the results for the binomial test for Model 3(B) show that the coefficient is not significantly negative for the majority of the regressions (proportion negative=0.10). Overall, investor affluence seems to have little significance with respect to an individual's decision to purchase local utility stocks. If affluent investors make better investment decisions (possibly because they have more resources) than regular investors, then my finding casts some doubt on the possibility of the local utility investment being driven by superior information about the stocks.

An investor's financial sophistication might influence his/her investment decisions. If sophisticated investors showed a greater tendency to hold utility stocks, then this might suggest an information based explanation for why individuals prefer their local utility stocks. To proxy for investor sophistication I use diversification measures in Model 4 & 5. Affluent investors in my data were generally more diversified then their non-affluent counterparts.[23] To avoid problems of multi-collinearity I excluded the Affluent Investor Dummy in Model 4 & 5. Model 4 uses the average number of stocks in an investor's portfolio over the 71 month period as a measure of investor sophistication. The results in Table 3.2, Panels B & C indicate that investors holding a greater number of stocks tend to purchase their local utilities. However, the binomial test (p-value=0.47) showed that this conclusion is a weak one. Investors could be naively diversifying by ignoring correlations among multiple securities. Thus, I use average correlations among securities as an alternative measure of diversification in Model 5. A mean coefficient of -0.78 for average correlations might suggest a role for investor sophistication. However, the proportion of negative coefficients in Table 3.2, Panel D, is only a significant 0.17. Thus there appears to be little impact of diversification ability on an individual's impact to purchase their local utility stock.

For all models, the Local Dummy is still significant after controlling for various factors. The odds ratio based upon the average and median $\beta_1$ (coefficient for Local Dummy) estimates range from 3.82-4.66 and 4.10-5.16, respectively. These results are validated by the high level of significance observed in the binomial test. Prior performance of stock, investor affluence, and portfolio diversification measures have little relation to an individual's decision to purchase utility stocks. Overall, an individual preference for local stocks seems to be a dominant factor influencing an household's decision to select among various utility stocks.

[23] The mean number of stocks in affluent and non-affluent investor portfolios is 7.39 and 5.45, respectively. The mean of the average correlations among securities in an individual's portfolio for affluent and non-affluent investors groups is 0.38 and 0.42, respectively. The differences between affluent and non-affluent investors groups are significant at the 1 percent level for both the measures of diversification measures. I draw similar conclusions from other diversification measures used in Goetzmann and Kumar (2008). Affluent investors as a group seem more diversified portfolios in comparison to the non-affluent investors.

## 6. Explanations for individuals preferring their local utility stocks

An investor's preference for his local utility could be attributed to one of the following: (1) information based explanation, (2) hedging utility expenditure, and (3) familiarity bias. The information based explanation refers to individuals possessing superior or private information about their local utility because of geographical proximity to the company. The hedging argument refers to individuals using their investments in local utility companies to hedge against possible increases in their utility costs. Familiarity bias refers to investors investing in their local utility companies simply because they feel comfortable investing in a familiar stock.

### 6.1. Information Based Explanation

An informational advantage for local investors may derive from substantial coverage of local firms by the local press and media, interaction with employees, executives and other parties involved in business transactions with a firm. Coval and Moskowitz (1999) find that US investment managers exhibit a strong preference for locally headquartered firms, particularly small, highly levered firms that produce nontraded goods. Analyses of the same large discount brokerage dataset used in my study, but including non-utility stocks as well, Ivkovis and Weisbenner (2005) attribute individual preference for local stocks to the presence of information.

The presence of any informational advantage in local utility transactions should be born out through positive abnormal returns in local utility roundtrip trades. This section explores the performance of local utility portfolios of individuals in the various contexts to check for the presence of any information.

### 6.1.1. Local Vs Non-Local Utility Stocks

The presence of superior or private information should lead to positive abnormal returns on roundtrip trades in local utility stocks. Furthermore, if information privileges are attributable to geographical proximity, then local utility roundtrip trades should outperform the non-local utility roundtrip trades. Using various abnormal return measures, I explore the average roundtrip of utility trades at both the trade and household level in Table 3.3, Panel A. For both the trade and household level, no significant abnormal returns accrue to local utility stock trades. Using the Carhart (1997) four factor model, the monthly abnormal return earned by the average household on local and non-local utility stocks was 0.18 (p-value=0.19) and 1.29 (p-value=0.30) percent, respectively. Contrary to the implications of the information based explanation, non-local utility stocks seem to outperform local utility stocks. However, the differences remain statistically insignificant.

In Table 3.3, Panel B, I explore the differences in gross returns between the following types of households: (a) Households that purchase only local or non-local utility stocks, (b) Households that purchase both local and non-local utility stocks. Purely local utility investors earn a monthly 4 factor abnormal return of 0.27 (p-value=0.21) in comparison to 1.59 (p-value=0.30) percent earned by purely non-utility investors. I find no evidence that households that strictly adhere to purchasing their local utility stocks earn significant positive abnormal returns or outperform households that buy only their non-local utility stocks. Among households that hold both local and non-local utility stocks, the differences between the local utility and non-utility trades remain insignificant. Overall, the results indicate that possible information asymmetry fails to explain why individuals are so likely to invest in their local utility stocks.

### 6.1.2. Non-S&P 500 Vs S&P 500 Local Utility Stocks

Among utility firms, opportunities to profit from information are likely to be greater for firms not listed in the S&P 500 index as these stocks are likely to receive little attention from analysts and the national media. The presence of superior or private information requires the following: (1) Individuals should earn significantly positive abnormal returns on their Non-S&P 500 local utility stocks, and (2) The realized abnormal returns on Non-S&P 500 local utility stocks should be significantly greater than their S&P 500 local utility stocks. I find that the abnormal returns for both the Non-S&P 500 and S&P 500 local utility stocks, (refer Table 3.4 Panel A) are statistically insignificant at both the trade and household level, indicating a lack of private or superior information about either kind of stock. These results hold in Table 3.4, Panel B, which looks at various household classifications. Interestingly, the Non-S&P 500 utility stocks seem to outperform the S&P 500 utility stocks, hinting at the possibility of some information advantage. However, the differences in abnormal returns between Non-S&P 500 and S&P 500 utility stocks are statistically insignificant.

### 6.1.3. Regional Vs Non-regional Local Utility Stocks

Another possibility is that local investors are likely to benefit more from investing in utility firms operating in only one state (referred to as regional utility stocks) as opposed to firms operating in multiple states (referred to as Non-regional utility stocks). This relates to possible mispricing of stocks with limited information dissemination. The results in Table 3.5 are similar to the analysis presented in the previous section. The abnormal returns are statistically insignificant for both Region and Non-regional local utility investments. The performance of regional local stocks tends to be better than the Non-regional local stocks, but the difference are statistically insignificant.

#### 6.1.4. Robustness check

To check the robustness of my previous results to outliers, I winsorized the top and bottom 5 percent of monthly returns for all utility roundtrip trades. The results are not presented in the interest of conserving space. I find that the abnormal returns at the household level for non-local and local utility stocks are negative and the differences are statistically insignificant.

Revisiting the performance of Non-S&P 500 and S&P 500 local utility stocks with the winsorized data, I find that investors earn negative abnormal returns on both types of stocks. The differences between Non-S&P 500 and S&P 500 stocks are mostly statistically insignificant. Even the few cases when they are statistically significant based on some measures of abnormal returns, those differences are miniscule. For example, based on the Fama-French 3 factor Model, the difference between Non-S&P 500 and S&P 500 local utility roundtrip trades is 0.29 basis points per month, which is economically insignificant.

The comparison of Regional and Non-Regional local utility stocks mirror the results based on the Non-S&P 500 and S&P 500 classifications. After a thorough study of realized monthly roundtrip utility trades, I fail to accept an information based explanation for individuals investing in their local utility stocks.

### 6.2. Hedging Explanation

Individuals could be investing in their local utility stock to hedge themselves against unexpected increases in their utility bills. An increase in utility prices results in an increase in monthly utility expenditure, which decreases the consumption bundle available for a household. Hedging could be an effective argument if an increase in utility prices results in an improvement in profitability for the utility provider, which could positively impact returns on the company's stock. In other words, holding a long position in local utility stock is consistent with hedging if there is a positive correlation between the movement in utility prices and returns on the local utility stock.

The data for this study covers 177 direct utility firms, with the majority of firms involved in supplying electricity (107 firms) and gas (50 firms). Testing the validity of hedging requires information on monthly prices for utilities in each state. Due to data constraints I restrict my tests to electric and gas companies. The data on electricity and natural gas prices for residential customers in each state from Jan 1991 to Nov 1996 was obtained from the U.S. Department of Energy, Energy Information Administration website. To test the validity of using local utility stocks to hedge one's monthly utility expenditure, I test to see if there is a positive correlation between the percentage change in electricity or gas prices in each state and returns on a portfolio of utility stocks that operate within that state during the 71 month period. For each state, I calculate the correlation between the monthly percentage change in energy prices and the monthly return on a portfolio of companies operating in that state. Both equal and value weighting specifications are used for measuring each state's utility portfolio returns. Table 3.6 presents a summary of the following correlations across all the states: (1) percentage change in utility prices and equal-weighted returns, and (2) percentage change in utility prices and value-weighted returns.

For electricity utility firms, the mean correlation between the percentage changes in prices and equally weighted returns across 50 states in the United States is -0.0561 (significant at $\alpha=0.05$). These results are supported by mean and median correlations between prices and value weighted returns. Contrary to expectation, the mean and median negative correlations across all states indicate that hedging against a rise in a household's electricity bills generally requires holding a short position in local direct electricity stocks. Virtually all households in my sample hold long positions in their local direct utility stocks, making the hedging argument ineffective.

For gas utility firms, the mean and median correlations are slightly positive but are not significantly different from zero. The zero correlation for gas utility stocks indicates that any position in local gas utility stocks would be ineffective for the purpose of hedging against a possible increase in natural gas price.

Overall, I find little evidence to support the argument that a long position in local direct utility stocks could be consistent with households hedging themselves from an unexpected increase in utility prices. One reason for the ineffectiveness of these hedges could be the heavy government regulation of utility firms. Massa and Siminov (2006) in their study of Swedish investors find that individuals do not engage in hedging, but invest in stocks closely related to their non-financial income. From a diversification perspective, Huberman (2001) states that a person should diversify and invest less in local companies serving him than in those operating in other parts of the country because the fortune of the local companies varies with economic tides in their areas.

### 6.3. Familiarity Bias

In the previous sections I reject the information and hedging explanations to individuals' bias for investing in their local utilities. This leaves us with the pure familiarity based explanation. Faced with scarce information about their non-local utilities (i.e., out of state utilities), individual seem to feel comfortable investing in their familiar local utility providers. This conclusion is not surprising as individuals often use behavioral heuristics to simplify their investment decisions. This result complements the finding in Huberman (2001), that familiarity drives individuals to over-invest in their regional (in-state) bell company. Grinblatt and Keloharju (2001), in their study of the Finnish stock market, find that investors simultaneously exhibit a preference for nearby firms and for same-language and same-culture firms.

## 7. Are local utility investors naïve?

One could argue that investors who purchase their local utilities are naïve, making the study of their behavior an inconsequential exercise. Table 3.7 presents various measure of diversification and compares them across the following categories of individual investors: (1) Non-utility stock investors, (2) Non-local utility investors (i.e. utility investors that never bought their local utility stock), and (3) Local utility investors (i.e. utility investors that bought their local utility stock at least once). Using five different measures of diversification,[24] I find that utility investors are more

---

[24] The first measure used is the number of stocks (STK). Although the STK measure does not account for correlation among securities, it is the simplest measure and has been used in several other studies (Blume, Crockett and Friend (1974) and Vissing-Jorgensen (1999)). A higher STK is likely to indicate greater diversification. The second measure considered is the sum of weighted squares (WSQ) used in Blume and Friend (1975). It is calculated as:

$$\text{WSQ} = \sum_{i=1}^{N} \left(w_i - w_{im}\right)^2 \approx \sum_{i=1}^{N} w_i^2$$

where $w_i$ is the proportion invested by a household in security $i$, $w_{im}$ is the proportion of security $i$ in the market portfolio and $N$ is the number of securities. WSQ attempts to determine how closely a household replicates the market portfolio and is measured by summing the squared deviation of the proportions invested in each security from the proportion in the market portfolio. Since the proportion in the market portfolio is small, this measure can be approximated as the sum of the square of the proportion invested in each security. A lower WSQ measure indicates greater diversification.

The third measure of diversification is the number of industries across which households diversify their holdings (IND). All stocks are classified across 10 broad based industry classifications, obtained from Kenneth French's website. Although this measure does not capture the exact correlation among stocks, it is a better indicator of diversification than the STK measure. A higher IND measure indicates greater diversification.

The fourth measure considered is the industry weight square (IWSQ) measure and is calculated as:

$$\text{IWSQ} = \sum_{ind=1}^{N} w_{ind}^2$$

where $w_{ind}$ is the proportion invested in each industry and $N$ is the number of industries. IWSQ measures the concentration of investments across households and is similar to the Herfindahl-Hirschman Index (HHI), a diversified than non-utility investors. Utility firms are a small portion of the market capitalization of the US stock market and thus one would associate investments in these firms with investors who maintain more diversified portfolios. In comparison to the average 5.32 stocks held by non-local utility investors and the average 6.03 stocks held by local utility investors, non-utility investors on an average hold 2.85 stocks in their portfolio. Interestingly, based upon measures like the number of stocks, the weight square and the number of industries owned, I find that individual investors buying their local utilities tend to be more diversified than those that buy only non-local utilities. Though these results are statistically significant, the differences in diversification are small among these two kinds of utility investors. The industry weight square measure and average correlation measures indicate that the differences in diversification between non-local utility and local utility investors are insignificant. I can safely conclude that individuals buying their local utility stocks are not naïve, as their understanding of portfolio diversification is as good, if not better than both non-utility and non-local utility investors.

## 8. Wealth effects and preference for local utility investments.

Could local utility investor's be the less-affluent investors who have access to resources that aid better financial decision making? The discount brokerage house classifies all investors who maintain an average balance of $100,000 as affluent. I find that the percentage of affluent investors among non-local and local utility investors to be 21.00 and 25.02 percent, respectively. The slightly greater proportion (p-value<0.01) of affluent investors among local utility investors does not support the argument that investing in a local stock is confined to people who do not

---

commonly accepted measure of concentration of firms within an industry. A lower IWSQ shows lower concentration of holdings within few industries, indicating a better diversification strategy.

The last measure is the average correlation among securities in an investor's portfolio.

have extensive access to investment resources. To further check the robustness of these results, I compared the self-reported net-worth for non-local and local utility investors and did not find any significant differences in wealth.[25]

## 9. Who makes money on their local utility stocks?

My results indicate that individuals fail to earn abnormal profits in their local utility investments. Their preference for local utilities seems to be driven by familiarity instead of some private/superior information. However, it is interesting to see if any specific segments of investors tend to make money on their local utility roundtrip trades. To compare performance of investors I use abnormal returns calculated using the Carhart (1997) 4 factor model.[26] Contrary to expectations, 23 percent of individuals that did not earn positive abnormal returns were affluent, while only 14 percent of individuals that earned abnormal returns were affluent. This finding suggests that being wealthy does not necessarily translate to access to superior information.

Furthermore, I classified local utility investors into categories based upon the number of average stocks in their portfolio, which is the simplest measure of diversification. I find that investors holding less than ten stocks, a category that includes the majority of investors in my data, make an abnormal monthly return insignificantly different from zero. Those who are well diversified make a statistically significant monthly abnormal return of 0.90 percent. Thus, it appears that for this small sample of investors, those who tend to be well diversified also make better investing decisions.

[25] Only 4,663 out of the 11,186 household analyzed, reported their net-worth.
[26] My results remain robust other factor models.

## 10. Conclusion

This study finds that individuals tend to prefer their local utility stocks as opposed to non-local utilities. My tests indicate that the information based explanation fails to explain this behavior. Also, I find no evidence that local utility stocks can be used to hedge a household's expenditure on utility costs. Thus, the preference for familiar assets, referred to as familiarity bias, is my explanation for this behavior. I find that local utility investors are as, if not more, affluent or sophisticated than their non-local utility investing counterparts. Thus, I find no support for the argument that individuals investing in their local utility stocks are naïve or less-affluent in comparison to other investors. The only individuals who seem to be earning positive abnormal returns on their utility stocks are highly diversified investors, holding more than ten stocks in their portfolios, which is a fairly small portion of the investors at the large brokerage house.

Faced with large information asymmetries surrounding utility firms and cognitive resource limitations to learning about many firms, individuals prefer their familiar local utility stocks. Moreover, investing in local utility stocks seems to be fairly pervasive behavior, not confined to a particular class of investors.

# CHAPTER 4

# INDIVIDUALS AND OTC (OVER-THE-COUNTER) STOCKS: PREFERENCE FOR THE UNFAMILIAR

**Abstract**

*Using a discount brokerage house data, this study explores commonly held beliefs about individuals investing in Over-the-counter (OTC) stocks, presumably unfamiliar assets. Contrary to popular perceptions associated with gambler or lottery buyers, I find that investors are older, wealthier and more experienced at investing than their counterparts. Individuals investing in OTC stocks display a greater degree of diversification and have large portfolio turnovers.*

## 1 Introduction

The financial press extensively covers stocks listed on the NYSE, NASDAQ and AMEX. But little do we hear about stocks solely traded on the over-the-counter market, with the exception of 'broiler room' stories and 'pump and dump' schemes. What little attention that is received is generally negative in nature. Interestingly, individuals still invest in over-the-counter (OTC) stocks. Most of these stocks are listed on the OTC Bulletin Board (OTCBB) and Pink Sheets, which have been traditionally known as a heaven for penny stocks. Scholars have investigated over-the-counter stocks and penny stocks to an even lesser degree.

A significant contribution of this study is that using Odean's (1998) discount brokerage house data for the period January 1991 to November 1996, it tests some commonly held beliefs about individuals that invest in OTC stocks. OTC stocks form a small portion of the market capitalization of all stocks in the US stock market and the common perception is to associate such stocks with demographics that might be common with lottery buyers/gamblers/speculators. The National Gambling Impact Study (1999) found that low-income individuals participate in lotteries at a much higher rate than do higher-income players. Thus it is plausible to assume OTC investors to be poorer, inexperienced and less sophisticated investors. Contrary to expectation, I find that these assumptions are mistaken. This study explores OTC investors and their association with demographic variables such as Gender, Age, Wealth (Net-worth), Investing Experience, Marital Status and Retirement Status. The portfolio characteristics studied are: (1) Measures of Non-OTC portfolio diversification, (2) Non-OTC portfolio turnover and (3) Proportions of the portfolio invested in large cap and micro cap securities.

If OTC investors were to be under-diversified relative to others, then their preference for firm-specific risk is evidence of risk seeking behavior. However, greater diversification for OTC stocks is consistent with Shefrin and Statman's (2000) behavioral portfolio theory, wherein investors choose portfolios by considering expected wealth, desire for security and potential, aspiration levels, and probabilities of achieving aspiration levels. Investors forming behavioral portfolios tend to divide their portfolio into layers and within each layer they have different attitudes towards risk. With a small portion of their portfolio they might be willing to take a huge amount of risk by investing in instruments such as lottery tickets and potentially highly volatile OTC stocks for a shot at big gains. Where else could one make an astonishing 1400 percent return within a day (Refer to statistics for "top gainers (%)" in table 4.1, Panel D)? Investors investing in an unfamiliar asset class such as OTC stocks are probably active traders, which leads me to explore their portfolio turnover. Also, investors holding a greater proportion of large cap stocks display a stronger preference for visible firms are unlikely to be OTC investors. In addition, this study explores asset allocation decisions of OTC investors, market conditions around OTC trade execution and performance of OTC roundtrip (from buy to sell trade) trades

Although investment in OTC stocks forms a small portion of an average households' portfolio (2.69 percent), OTC stock ownership is pervasive. I find that nearly half of the (32,108 out of 65,591) households that invested in stocks at the discount brokerage held a position in an OTC stock. This study finds that contrary to characteristics associated with lottery buyers/gamblers/speculators, individuals who are older, relatively wealthier, and experienced at investing are more likely to invest in OTC stocks. The observed results can not be attributed to individuals using play money accounts that account for an insignificant portion of their total wealth. Interestingly, individuals who are more diversified in their Non-OTC portfolios (lower portfolio turnover and relatively more diversified portfolio) are more likely to be OTC stock investors. Based on observed returns, I find a tremendous downside risk in OTC stocks priced under $100, which is compensated by the low probability of extremely high returns. One must be cautious not to label those investing in OTC-stocks as naïve investors looking to gamble in the stock market. Indeed, most of these investors tend to be more sophisticated (affluent, experienced and relatively more diversified in their stock portfolio) individuals who invest their wealth in layers of risky and safer assets, an implication of Shefrin and Statman's (2000) theory of behavioral portfolios.

## 2 Literature review and beliefs about Over-the-Counter (OTC) Stock

### 2.1 Literature Review

The literature investigating penny stocks or describing the general trading activity of investors who invest in them is sparse. This can be primarily attributed to the lack of data available in widely used financial databases like CRSP and limited access to household trading data. Hanke and Hauser (2008) in their study of the effects of stock spam emails on prices of over-the-counter securities (Pink Sheets and OTC Bulletin Board stocks) find that positive news contained in stock spam emails had no lasting positive effect on stock prices. Apart from Hanke and Hauser (2008), the little attention penny stocks have received in the financial literature covers delisting consequences from exchanges and penny stock IPOs. Harris, Panchapagesan and Werner (2008) find that firms that are delisted from NASDAQ and are relegated to the OTC Bulletin Board and Pink Sheets experience a large decline in liquidity, which is also associated with a significant decline in wealth. Bradley, Cooney, Dolvin and Jordan (2006) in their study of penny stock IPOs, find that they have higher initial returns than ordinary IPOs, but significantly worse long-run underperformance. Due to a lack of data, their study did not include IPOs that initially started trading on the OTCBB, Pink Sheets or the grey market. They do include offerings that initially started trading on the Nasdaq SmallCap Market with an offer price of less than or equal to $5. Beatty and Kadiyala (2003) find that the Penny Stock Reform Act of 1990 (PSRA) had the cosmetic effect of reducing the number of IPOs priced below $5 but had no substantive impact on issuer quality.

### 2.2 Over-the-Counter (OTC) Stock Description

Over the counter (OTC) stocks refer to securities that are traded either on the OTC Bulletin Board or on Pink Sheets, or both. These stocks are generally considered to be highly speculative and risky because of their lack of liquidity, large bid-ask spreads, small market capitalization and limited following and disclosure. OTC stocks are subject to limited listing requirements along with fewer filing and regulatory standards. OTC stocks can also be called penny stocks. The term penny stock is itself a misnomer because there is no generally accepted definition of a penny stock. Some consider it to be any stock that trades for pennies or those that trade for under $5, while others consider any stock trading off of the major exchanges (NYSE, NASDAQ or AMEX) as a penny stock. The US Securities Exchange Commission (SEC) states: "The term 'penny stock' generally refers to low-priced (below $5), speculative securities of very small companies. While penny stocks generally trade over-the-counter, such as on the Over The Counter Bulletin Board (OTCBB) or in the Pink Sheets, they may also trade on securities exchanges, including foreign securities exchanges. In addition, penny stocks include the securities of certain private companies with no active trading market." For the purpose of my study, OTC stocks will be considered synonymous with penny stocks.

Trading in OTC stocks is often sporadic and erratic. This has three negative impacts. First, thin trading ensures that a few market makers can manipulate prices by controlling the market (Chen, 2002). Second, the infrequent trading in penny stocks may make them difficult to sell (SEC). Lastly, it may be difficult to find quotations for certain penny stocks, thus making it difficult to accurately price (SEC).

There is no listing fee for securities on either the OTCBB or the Pink Sheets. The filing requirements are slightly higher for OTCBB as compared to Pink Sheets. Issuers of all securities quoted on the OTCBB are subject to periodic reporting of financial information to the SEC, banking, or insurance regulators. However, McLean (2000) mentions that historically half the issuers have failed to do so. The SEC stigmatizes some companies with an extra "e" at the end of their ticker to indicate non-compliance with filing requirements and might even kick some companies off the bulletin board. On the other hand, issuers on Pink Sheets are not required to register securities with the Securities and Exchange Commission (SEC), or be current in their reporting requirements to be quoted. Nor are they required to file financial or other company information with the Pink Sheets. In order to protect and warn investors, Pink Sheets places small icons before each security's ticker symbols to indicate the company's regularity in filing reports with regulatory organizations, but does not impose any reporting standards. Overall the listing requirements for all OTC stocks are minimal. Also, OTC stocks are hardly ever followed by financial analysts, adding to the information asymmetry.

The absence of rigorous listing requirements as well as the high possibility of price manipulation makes OTC stocks an easy target for fraudulent activities. Also the presence of little information makes these securities highly speculative in nature. The SEC warns: "Investors in penny stocks should be prepared for the possibility that they may lose their whole investment."

However, it would be cynical to view OTC stocks as junk. The financial press is replete with stories about fraud and scams in the penny stock market, but few talk about some well run companies that might have untapped invest value. Roane (2007) states that the problem the CEO of the Pink Sheets, R. Cromwell Coulson, faces is that well-publicized stock manipulations on Pink Sheets obscure the fact that there are well-run and profitable companies listed there, often trading at large discounts to exchange-listed rivals. There are solid community banks that already follow stringent federal reporting requirements, large overseas companies like Nestlé and Wal-Mart de México that are comfortable with being listed only on their home exchanges and promising small companies with real revenues that can't justify the regulatory cost of listing on a true exchange. In addition, large financial institutions like UBS Securities LLC, Citigroup Capital Markets, and Merrill Lynch, participate in these markets as market makers.

### 2.3 Why Invest in OTC stocks?

Hanson and Richards (2006) find that typing in the term “penny stocks” to the Google search engine produced “about 1,210,000 hits”. When they did the same for more time-tested terms such as “blue-chip stocks” and “dividend-paying stocks” they found 266,000 and 173,000 hits, respectively. This measure has its share of discrepancies, but nevertheless, it does indicate an element of investor interest in penny stocks. Hanson and Richards (2006) state: “We love penny stocks because they’re fascinating. The world of pennies is inhabited by hardworking average Joes hoping to strike it rich, pumpers and dumpers, hypesters and scammers. In pennies, the logic and reason that applies in the rest of daily life is replaced by zeal and prayer.” For many investors playing in OTC stocks is akin to playing the slot machines. Many investors are aware that they are likely to loose their entire investment. Roane (2007) also attributes interest in OTC stocks to the lure of quick gains on a long shot. Kumar (2008) finds that the individual investor’s demand for lottery type stocks (stocks with high variance, high positive skewness and low prices) increases when times are poor and these demand shifts influence the returns of lottery type stocks.

OTC stocks could lure investors who sometimes perceive them to be under-valued and neglected stocks. During times of optimism, some investors are ready to take a chance on well run companies that might not be traded on standard securities exchanges. McLean (2000) feels that gut attraction could be the result of the law of small numbers. It might seem easier to make huge percentage gains on a stock that costs 1 cent a share than on one that costs $100. James Angel, in an article (Roane, 2007) about trading on Pink Sheets says that it’s “like the kiddie pool at the swim club. It’s too small for the adults, but the little people have a great time.”

### 2.4 OTC stocks and speculative forces

OTC stocks have a reputation for being speculative in nature. McLean (2000) describes the penny stock boom prior to the technology sector bubble burst and mentions two semi-conductor stocks amongst many that experienced tremendous returns: Illinois Superconductor and International Superconductor. In two short months, from Dec 29, 1999 through Feb 28, 2000, Illinois Superconductor shot up over 7,500%. This was nothing compared to International Superconductors' returns of 32,400% over just 18 days. During that time period, most investors were optimistic, maybe irrationally exuberant, about technology stocks that had fundamentally lop-sided stock valuations. The excesses were committed not only on the part of individual investors, but most institutional investors as well. McLean (2000) states that bulletin board may be one of the best indicators I have of just how deep and wide the speculative fever infecting the market really is. In other words, activity in penny stock activity could be a good indicator of market sentiment.

OTC stocks tend to be shunned by institutions. However, many individual investors who do not normally trade penny stocks may get drawn into them during periods of high sentiment. Nevertheless, penny stock activity may capture the most speculative nature of sentiment. Highly speculative fever of some individual investors may lead to a greater sentiment of both other individual investors and eventually institutions.

## 3 Data and Methodology

### 3.1 Data

This study uses the dataset of trading activity of individuals at a discount brokerage house during the period January 1991 to November 1996. Each security is identified by the brokerage house with its CUSIP number. The CRSP database consists only of stocks listed on the NYSE, NASDAQ and AMEX. By definition, I classify all securities that are not listed on the NYSE, NASDAQ or AMEX trade as OTC/penny stocks. Thus if I am unable to find a CUSIP match for a security on CRSP, I identify it as an OTC stock. Since listing on Pink Sheets or OTCBB does not require a CUSIP number, securities with no CUSIP identifiers are also identified as OTC stocks.[27] In the absence of detailed data for OTC stocks, prices for OTC stocks are inferred through the trades and holdings of the Odean (1998) stockbroker dataset. Approximately 35% of the households that traded in stocks traded in at least one OTC stock during the 71 month period.

Table 4.1 provides a summary of 75,942 trades made by 21,782 households in 5,307 OTC stocks. Panel A presents the distribution of the average price of each penny stock traded at the discount brokerage house. The median security price is barely $0.675 (mean of $8.47). However it seems that the majority of the trading occurs in higher priced stocks. Panel B shows that the median purchase trade has a price of $7.625 (mean of $44.48). Some of the trades have large dollar values indicating trades possibly made in privately-held companies. The dataset also includes month end positions and closing prices. I have identified 2,500 additional securities that appear in the position statements, but not in the transaction statements. This can be attributed to the transfer of securities that might have been made from or to their accounts at other brokerage houses. Such transfers are not noted in the transactions statements as they are not actual trades. In addition, some of the penny stocks may have been purchased before the start of the available data.

[27] The brokerage house also has its own security number for each stock traded. However this security number could vary for a stock with a change in CUSIP. In the absence of a CUSIP identifier, we use this security number to identify OTC stocks.

In Table 4.1, Panel C, I present the market statistics for the trading activity at the OTCBB exchange on March 8, 2009, a randomly chosen date.[28] The prices in the table range from $.0001 to $12,500, which lends credibility to the observed numbers in the discount brokerage sample. An interesting point to note is that even in the ensuing financial crisis that started in 2007, the top gainer made 1400 percent returns while the biggest looser made -88 percent returns at the OTCBB.

The brokerage data also includes a file (the Infobase data) that includes demographic information like age, gender, marital status, retirement status and income of the account's owner. Although this demographic data is only available for a subset of the total brokerage accounts, it is essential to understanding the household characteristics that are associated with investing in OTC stocks.

## 3.2 Methodology

### 3.2.1 Formulation of Price Series Using Adjustment Factor

For all non-CRSP stocks, I have built a time series of adjustment factors for each security using the transaction prices and end of month prices obtained from position and transaction statements of households. I determine a price adjustment factor for each security based on the number of shares held by a household at the beginning of each month, the number of shares traded during the month and the number of shares held by the household at the end of the month. For example, assume that a household has 100 shares of stock XYZ on 1 Jan, 1991. It trades no shares during the month of January and has a closing position of 200 shares in that stock. This indicates that either a stock dividend has been paid to the share holders or a possible stock split has occurred. I will use this to adjust future prices of shares of XYZ. One potential concern with this

[28] The table presented was downloaded from the OTCBB website (www.otcbb.com) on March 8, 2009 at 7:40pm.

methodology is that at times, households transfer shares from another brokerage house and these are not reflected in their transaction statements. Adjustment factors can be verified for some stocks that are held by multiple households during the same month, but the others remain unconfirmed. Using this procedure, I have adjusted the prices of 225 securities, of which 181 security price adjustments remain unconfirmed.

### 3.2.2 Return Calculations

Due to the lack of information about cash dividends paid by penny stocks, I will be able to calculate returns solely from capital appreciation/depreciation. It seems unlikely that penny stocks pay regular cash dividends and thus returns from capital appreciation should be fairly close to the overall returns on most stocks.

The gross return for a roundtrip transaction on security i is calculated as:

$$R^{gr}_{i,t+1} = \frac{P^{s}_{i,t+1} - P^{b}_{i,t}}{P^{b}_{i,t}}$$

where $P^{s}_{i,t+1}$ is the adjusted sale price of at time t+1, and $P^{b}_{i,t}$ is the adjusted purchase price at time t. The formulation of the adjusted price series is discussed in section 5.1. The net return for a roundtrip transaction on security i is computed by adjusting the gross returns for commission costs as:

$$R^{net}_{i,t+1} = \frac{P^{s}_{i,t+1}(1 - c^{s}_{i,t+1}) - P^{b}_{i,t}/(1 + c^{b}_{i,t})}{P^{b}_{i,t}(1 + c^{b}_{i,t})}$$

where $c^{s}_{i,t+1}$ is the cost of sale scaled by the price ($P^{s}_{i,t+1}$) at time t+1 and $c^{b}_{i,t}$ is the cost of purchase scaled by the price ($P^{b}_{i,t}$) at time t.

The gross returns for all OTC trades made by a household h during the period January, 1991 to November 1996 is calculated as

$$R_h^{gr} = \sum_{j=1}^{n_j} w_j R_j^{gr}$$

where $R^{gr}_j$ refers to the gross returns on round trip transaction j (i.e. $R^{gr}_{i,t+1}$), and $w_j$ refers to the dollar value of transaction j scaled by the dollar value of all repurchases made by household h. Similarly, the net return for all repurchases made by household h is calculated as

$$R_h^{net} = \sum_{j=1}^{n_j} w_j R_j^{net}$$

where $R^{net}_j$ refers to the gross returns on round trip transaction j (or $R^{net}_{i,t+1}$), and $w_j$ refers to the dollar value of transaction j scaled by the dollar value of all repurchases made by household h. The gross and net returns for all repurchasing households assigned to a particular category are calculated as:

$$R_k^{gr} = \frac{1}{n_k} \sum_{h=1}^{n_k} R_h^{gr} \quad and \quad R_k^{net} = \frac{1}{n_k} \sum_{h=1}^{n_k} R_h^{net}$$

where $n_k$ is the number of households in a category k.

### 3.2.3 Missing demographic variables

While conducting multivariate regressions which involve demographic data, one econometric issue faced is the large number of missing variables that could lead to significant loss of observations and result in biased parameter estimates. This study uses two approaches for dealing with missing data in regression analysis: (1) Missing Dummy Approach, and (2) Multiple Imputation Approach. The missing dummy approach uses a dummy variable to substitute for missing valued but has been criticized for possibly yielding biased estimates. The multiple imputation technique is more sophisticated as it makes complex data imputations for missing values. I use one of the more advanced techniques of multiple imputations, the Markov

Chain Monte Carlo (MCMC) method, which assumes that the data is missing at random (MAR). The MCMC approach generates a stationary distribution for the variables in the regression equation and using a series of steps imputes the missing values. A detailed discussion is presented in Schafer (1997). The estimates for each of the regressions for each complete imputation, *m*, is calculated as follows:

$$\text{Logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_{0,m} + \sum_{j=1}^{n} \beta_{j,m} X_{j,i}$$

Rubin (1987) shows that 5 complete data imputations are sufficient to achieve a reasonably high level of efficiency. The estimated coefficients based upon the 5 dataset imputations are calculated as follows:[29]

$$\beta_j = \sum_{m=1}^{5} \beta_{j,m} \Big/ 5 \quad \text{for } j = 0,1,\ldots,n.$$

Consistent results between the two approaches (missing dummy and multiple imputations) lead to confidence in the regression estimates.

### 3.2.4 OTC quintiles

The behavior of households investing in OTC stocks is likely to differ based upon the portion of portfolio invested in OTC stocks. To account for any possible variations in OTC investor characteristics and preferences, I divide households investing in OTC stocks into quintile based on the ranking of the portion of their average portfolio invested in OTC stocks over the months where they held a position in at least one stock.

[29] Please refer to Rubin (1987) for further details.

# 4 Demographics and OTC Investors

## 4.1 Univariate tests

This section explores the demographic characteristics of OTC investors. Kumar (2009) finds that socio-economic and psychological factors which induce higher expenditures in lotteries also induce greater investment in lottery-type stocks (but still exchange listed) – poor, young men who live in urban, Republican dominated regions and belong to specific minority (African-American and Hispanic) and religious (Catholic) groups invest more in stocks with lottery-type features. If investing in OTC stocks was restricted to only to people who were pure gamblers or speculators then I would expect them to be poor young men with little investing experience. Table 4.2 presents univariate tests of the difference in demographic characteristics between households that never invested in a OTC stock (referred to as Non-OTC investors) and those who did (referred to as OTC investors).

Barber and Odean (2001) find that consistent with the prediction of over-confidence models, the average portfolio turnover for men is one and a half times that of women. There is also some evidence in support of the view that women are more risk-averse than men in financial decision making (Jianakoplos, Ammon and Alexandra, 1998, Barsky et al., 1997). Table 4.2, Panel A shows that 46.92 percent of males own OTC stocks, which is slightly higher than 41.96 percent for females. The results are statistically significant (p-value<.0001) and are consistent with the risk aversion argument.

Are OTC-stock investors simply naïve or inexperienced at investing? To answer this question I compare experience levels (high or low)[30] and OTC stock ownership (Table 4.2, Panel B). Contrary to the hypothesis of naïve investing, I find households with a high level investing

[30] Investors self reported their level of investing experience as either, Extensive, Good, Limited or None. To filter out some noise I have classified their response as follows (1) High experience category – those with Extensive or Good experience, and (2) Low experience category - those with Limited or no experience.

experience are more likely (p-value<.0001) to invest in OTC stocks than those with a lower level of experience. Approximately 54 percent of high experience households bought OTC stocks in comparison to 44 percent for the low experience households. Similar to testing for investing experience, I tested for investor knowledge levels and found that investors who consider themselves more knowledgeable are also more likely to invest in OTC-stocks. However, this result was expected as the investor response for knowledge and experience levels are highly-correlated. These additional results have not been presented to conserve space.

An individual's marital and retirement statuses are also likely to impact his/her risk aversion. Married individuals have greater responsibilities and should have higher risk aversion than singles. Also, retired individuals are more interested in capital preservation as opposed to growth, implying greater risk aversion relative to those who are employed. The test for marital status (refer Table 4.2, Panel C) does not provide strong support (p-value=.14) in favor of the risk aversion argument. For the retirement status (Table 4.2, Panel D) a p-value of .0004 does provide statistical support in favor of the risk aversions argument, though the difference is not economically significant.

In Table 4.2, Panel E, I test for the impact of age on OTC stock ownership. The household head's age is classified as Young (less than 40), Middle-aged (between 40 and 65), and Old (over 65). I find that the percentage of households investing in OTC-stocks is 44, 46 and 49 percent for young, middle and old categories, respectively. These differences are statistically significant (p-value<.0001) and cast doubt on the notion that investing in OTC stocks is attributable to risk-seeking younger individuals.

The National Gambling Impact Study (1999) found that low-income individuals participate in lotteries at a much higher rate than do higher-income players. This view is largely supported by existing literature (Clotfelter and Cook, 1989). If OTC investors are similar to gamblers then they should belong to the lowest wealth group among the households at the brokerage house. Also, one might argue that affluent investors are likely to have access to resources and advisors to make sound investment decisions. I use net-worth as a proxy for wealth and divide the households into four wealth groups[31]; (1) Q1: Net-worth less than $75,000 (2) Q2: Net-worth from $75,000 to $100,000 (3) Q3: Net-worth from $100,000 to $250,000 (4) Q4: Net-worth greater than $250,000. The percentage of households investing in OTC stocks monotonically increases (p-value<.0001) with the wealth quartiles. For the lowest quartile and the highest quartile, the percentage of OTC investors is 46.30 and 54.54 percent, respectively. The results imply that OTC-investors can not be assumed to be poor individuals taking a shot at the riches. Also, there is no evidence of affluent investors abstaining from investing in OTC-stocks.

The results in this study would be of interest if the equity positions at the discount brokerage house represent a sizable portion of an individual's wealth or else one could skeptically put aside the findings in this study as gambling observed in a play money account. I use the size to net worth ratio to estimate the portion of an individual's wealth invested through the discount brokerage house. Size to Net worth Ratio (SNW) is the ratio of the average monthly investment in common stock divided by the self reported net worth of the client. Interestingly, the average household investing in OTC securities has a SNW of 0.57, which is significantly greater than 0.33 for the non-OTC households. In other words, the OTC-investor accounts are not play money accounts, which makes these conclusions more relevant.

[31] I use quartiles to as they tend to provide a sufficient number of categories with equal group sizes. Any higher number of categories lead to unequal group sizes.

The univariate test results lead us to conclude that contrary to characteristics associated with lottery players or gamblers, I observe that OTC stock investors are older and relatively wealthier than their peers.

**4.2 Regression results**

Using the results from the above univariate tests, I proceed to perform a logistic regression modeling the decision (Yes or No) of households to invest in an OTC stock at least once during the 71 month period. The dependent variables consist of Male Dummy, Experience Dummy, Age Dummies (Young & Old), Net-worth (Wealth) Dummies, Married Dummy, Retired Dummy, and the size to net-worth ratio (SNW). The experience dummy is coded as high experience relative to low experience. For the age dummies, the young and old are coded with the middle age group as the reference category. The net-worth (wealth) dummies are coded with reference category being the lowest new-worth quartile (i.e. net worth less than $75,000). To obtain efficient estimates in the presence of missing values I implement the following two approaches that have explained in greater length in the methodology section: (1) Approach 1: Missing value dummy approach, and (2) Approach 2:Multiple Imputation approach.

The regression estimates are provided in Table 4.3. Note that the estimates based on both approaches have the same sign and similar magnitude, which lends credibility to my observed coefficient estimates. The coefficients for the regression conform to my findings in the previous section. I observe a monotonically increasing probability of investing in an OTC stock across the three age groups and wealth (net-worth) quartiles. While the young relative to the middle aged have an odds ratio of .92 for investing in OTC stocks, the older group has an odds ration of 1.09. The highest wealth quartile is 1.25 times more likely to buy an OTC stock than the lowest wealth quartile. Also, the significantly (p-value<.0001) positive estimates for the SNW variable in both the regression approaches, indicates that one cannot argue that individuals who invest in OTC - stocks are gambling in a play money account that represents an insignificant portion of their wealth.

### 4.3 Robustness (Alternate OTC investor classification) – Regression Analysis

One potential criticism of my earlier analysis relates to investor classification as OTC and Non-OTC investors. It is possible that some investors might not have originally intended to hold OTC stocks and they became owners of OTC stocks due to those stocks being delisted off the major exchanges. Ideally, these investors should be excluded from my sample of OTC investors. However, I do not have any identifiers available for OTC stocks to check their delisting status. This necessitates an alternate methodology in which I consider only the buy trades to determine whether an investor is an OTC or Non-OTC investor. An investor making a buy trade when a stock is traded in the OTC markets can be confidently labeled as an OTC investor, a conclusion that can not be accurately made for investors holding OTC securities that appear on the position statement but are not supported by a buy trade.[32] All other investors who conducted buy trades on stocks listed only on major exchanges (NYSE/NASDAQ/AMEX) are classified as Non-OTC investors. Using the regression methodology specified in section 4.2 with multiple approaches to dealing with missing data, I find that the results with the new OTC/Non-OTC investor classification presented in Table 4.4 are fairly similar. The experienced and wealthiest investors show a propensity to hold OTC stocks. The negative (-0.05) young age dummy coefficient indicates that the younger investors are less likely to hold OTC stocks relative to middle aged

[32] These securities could be carried forward from prior to January 1991 or transferred from some other brokerage account. It is impossible to state accurately if the investor bought these stocks when they were listed as OTC stocks or when they were listed on major exchanges and subsequently got delisted.

investors. The old investors do not show any statistically significant preference for OTC stocks different from that of middle aged investors. The Male dummy is no longer significant, implying absence of gender specific effects. Similar to my previous regression results, there is no evidence of marital or retirement status. The size to net worth ratio has a significantly positive coefficient indicating that the households owning OTC stocks invest a slightly greater portion of their wealth through the brokerage house.

Overall, based on the results of the individuals who tend to be highly experienced at investing, older and wealthier people have a higher probability of investing in OTC stocks relative to their peers.

## 5 Portfolio Characteristics and OTC Investors

This section studies the portfolio characteristics likely to be associated with OTC investors. The following portfolio characteristics are evaluated in this section: (1) Turnover measures, (2) Portfolio Diversification measures, and (3) Preference for large and micro cap stocks. For the purpose of this analysis I present results based upon the alternate robust OTC investor classification specified in section 4.3. This is because this classification not only more accurately captures OTC investors but also alleviates problems associated with the turnover measure.[33]

Table 4.5, Panel A presents statistics for portfolio turnover calculated using the methodology specified in Barber and Odean (2000).[34] The turnover is calculated separately for

[33] Some of the households never traded any stocks in the 71 month period. By using the robust classification I restrict our households to those who made at least one buy trade during the 71 month period thus making turnover more comparable. For example a household never traded any stock during the 71 month period and one of his securities got delisted to the OTCBB. This household could be labeled as an OTC investor in our "non-robust" classification and would be observed to have no turnover. Thus leading to biased results.

[34] Monthly purchase (sales) turnover is the beginning-of-month market value of shares purchased in month t-1 (or sold in month t) divided by the total beginning-of month market value of shares held in month t. The monthly turnover is defined as the average of purchase and sales turnover for each month. The average of the monthly

the non- OTC stock and OTC stock portfolios in this analysis. The results indicates that the average turnover for OTC stock investors is about 2 percent (p-value<.0001) more than the Non-OTC investors. This result is intuitive as those who adopt an active approach to managing their stock portfolios might consider investing in unfamiliar securities like OTC stocks. Market timing in OTC securities is important and passive strategies are unlikely to indulge in OTC stocks for this reason. For households that traded OTC stocks, a paired t-test of the average difference between their non- OTC and OTC portfolios revealed that turnover for their OTC stock portfolio was significantly (p-value <.0001) lower by about 3.11 percent. This finding is not surprising as OTC stocks are known to be thinly traded, which could possibly induce a lower turnover. Also, it is possible that some of the OTC stocks are investments made for a longer horizon.

Household portfolio diversification could give some insight into whether households are sophisticated versus risk seeking. In this study, I view a household as sophisticated if it shows significant abilities to diversify its investments in comparison to its peers. On the other hand, a household which is highly under-diversified could be perhaps regarded as a lottery-type investor or be risk seeking relative to others. If investing in OTC stocks was a reflection of being risk seeking then OTC-investors should be under-diversified relative to non-OTC investors. I use the following four measures of diversification: (1) number of stocks, (2) sum of squared weights of holding in each stock, (3) number of industries, and (4) sum of squared weights of holding in each industry.[35] The greater the number of stocks (or industries), the greater the diversification benefits a household is likely to observe. However, a higher weight square for stocks (or industries) indicates a greater concentration of portfolio in a few stocks (or industries), resulting

---

turnover over the number of months a household held a position in at least one stock is used to calculate the turnover for each household.

[35] The stocks are classified into 10 industries as per the industry classification definition of Kenneth French's website.

in reduced diversification benefits. I find that (refer to Table 4.5, Panel B) OTC investors hold an average of 5 stocks in their portfolio as opposed to 2 stocks held by their Non-OTC investors. The results indicate that OTC-investors are more diversified than their non-OTC counterparts both in term of the number of stock (or industries) and concentration of their holding. Thus, there is little evidence to support the notion that OTC investors are risk seeking or follow a naïve investing strategy.

Furthermore, to view the joint impact of various portfolio characteristics I perform logistic regressions modeling the decision to repurchase stocks.[36] The results are presented in Table 4.5, Panel C. Conforming to the results presented in Panels A and B of Table 4.5, I observe that a greater non-OTC stock turnover is associated with a lower probability of investing in OTC stocks, while greater diversification, proxied using the number of stocks, is associated with a greater probability of investing in OTC stocks. The regression for portfolio characteristics also includes proportions of portfolio invested in the following stock categories: (1) Large Cap, and (2) Micro cap.[37] The coefficient for large cap and micro cap securities are -0.09 and 0.02, respectively. The direction of these coefficients is not surprising. OTC stocks are likely to be unfamiliar to most investors. Perhaps those investors who strongly prefer large cap securities are less likely to buy OTC stocks. However, individuals who are adventurous enough to endeavor into micro cap securities would probably consider investing in OTC stocks. Micro-cap stocks are probably the closest to OTC stocks than any other stock traded in the US stock market.

[36] It must be noted I conduct these regressions without including demographic data as additional independent variables. This is because demographic variables are often closely related to portfolio characteristics. For example, Barber and Odean (2001) find that the average portfolio turnover rate for men is one and half times that of women. Goetzmann and Kumar (2008) study the impact of various demographics on household diversification measures.

[37] The definition I use are: (1)Large cap stock : market cap > $ 10 billion, and (2) Micro-cap stocks: market cap ≤ $ 300 million.

The findings for portfolio characteristics indicate that OTC-investors are more financially sophisticated, with lower portfolio turnover and greater diversification. This finding can be reconciled under the tenants of behavioral portfolio theory (Shefrin and Statman, 2000), which posits that investors construct their portfolios as layered pyramids, where the bottom layers are designed for downside protection, while top layers are designed for upside potential.

## 6 Asset Allocation and OTC Investors

In order to further test the implications of behavioral portfolios, I study the asset allocation of investor portfolios among equity and fixed income securities within various wealth categories. Equities consist of the following: (1) Stocks divided into large-cap, medium-cap, small-cap, micro-cap and over-the-counter stocks, (2) Equity Mutual Funds (includes Unit Investment Trusts).[38] Fixed Income Instruments are divided into: (1) Bonds and (2) Fixed Income Mutual Funds (includes Unit Investment Trusts). The average allocation to Bonds, Fixed Income Mutual Funds, Equity Mutual Funds, Large Stocks, Medium Stocks, Small Stocks, Micro Stock and OTC stocks for all households at the brokerage house is 6.68, 2.54, 18.22, 27.52, 14.95, 10.69, 13.47 and 2.69 percent (all have p-value <.01), respectively. Interestingly, Table 4.6 shows that the average household's portfolio mix irrespective of their wealth tends to be fairly similar with the average investment in OTC securities in the range of 2.16 to 3.15 percent. The results indicate that irrespective of their level of wealth, the average individual will allocate his/her money between layers ranging from safe to risky investments. Thus, there seems to be some evidence that, on average, individuals tend to divide their investment portfolios among various

[38] The definition for each stock category is: (1)Large-cap : market cap > \$ 10 billion, (2) Medium-cap : \$ 2 billion<market cap ≤ \$10 billion), (3) Small-cap : \$300 million<market cap ≤ \$2 billion, (4) Micro-cap : market cap ≤ \$ 300 million, and Over-the-counter stocks: OTCBB and Pink sheet stocks.

securities based on the perceived riskiness of each security category and not necessarily on a risk aversion dependent upon their level of wealth.

To study the investment habits of OTC and non-OTC investors, I study the distribution of their investments amongst various security classes. I find that the differences in allocation between OTC and non-OTC investors, though statistically significant, are fairly similar as shown in Table 4.6, Panel B. Similarly, the differences among the OTC quintiles (Table 4.6, Panel C), except for the highest quintile, are negligible. Thus, it is hard to argue that OTC investors differ from their non-OTC counterparts in their observed asset allocation at the discount brokerage house.

## 7 Conditions surrounding trades in OTC stocks

In this section, I investigate conditions under which individuals trade in OTC stocks. Figure 4.1(a) indicates that the individuals at the discount brokerage make their first purchase in OTC stocks when there is a run-up in the performance of the overall stock market, small capitalization stocks, and their non-OTC stock portfolio. The annualized return on small capitalization stocks over the past one, two, three and four months prior to the purchase of a penny stock stood at 70.89, 63.13, 11.9 and 6.32, respectively. The returns over one and two months prior to the first purchase are particularly high considering that during the period 1991-1996 the annual returns for the CRSP value-weighted index was 20.7 percent. Similar trends are observed for the market and household non-OTC stock portfolios. Analysis of all purchase trades (Figure 4.1(b)) and sales trades (Figure 4.1(c)) shows a similar run-up in the overall stock market, small capitalization stocks and their non-OTC portfolio returns. The evidence from purchase and sale trades indicates that individuals tend to time their entry and exit from OTC-stocks based upon the performance of the stock market as a whole and also upon performance of other small non-OTC stocks. I find evidence of trades in OTC stocks occurring during periods of high market sentiment, which provides some support for McLean's (2000) argument that penny stocks may be one of the best indicators of the depth of speculative fever infecting the markets.

## 8 Performance of OTC roundtrip trades across OTC quintiles

Propensity to invest in OTC stocks could also be related to an investor's stock picking abilities or some private information about OTC listed firms. This implies that individuals in the highest OTC quintiles are likely to outperform their lowest OTC quintiles. Table 4.7 shows that highest OTC quintiles investors earn gross nominal returns of 15.69 percent, which is greater than the 11.77 percent earned by the lowest OTC quintiles. However, OTC investors could be investing at different times for different lengths of time, making comparisons versus benchmarks more appropriate for judging performance. Interestingly both the returns on gross small-cap index adjusted returns and gross market adjusted returns were greater for the lowest quintile OTC stocks. The small-cap index adjusted gross returns for the lowest and highest OTC quintiles are -14.27 and -25.44 percent, respectively. Thus there is little evidence supporting the hypothesis that investors who invest a greater amount of their portfolio in OTC stocks are doing so because they have better stock picking abilities or some amount of private information. The results indicate that investors in any OTC quintile underperform the small-cap benchmark on average. In terms of market adjusted returns, the performance across quintiles is either insignificantly different from zero or is negative. Accounting for commission costs, the under-performance of OTC roundtrip trades is further accentuated. OTC securities are thinly traded and have higher commission percentages in comparison to securities listed on regular stock exchanges like

NYSE, NASDAQ and AMEX. Commission costs lowered performance of trades across various OTC quintiles in the range of 5.01 to 6.68 percent.

## 9 Performance of OTC roundtrip trades across price categories

McClean (2002) feels that the gut attraction for investing in OTC stocks could be their low price levels. It could be easier to make money on a stock that is priced at 1 cent versus a $100. To explore this hypothesis, I compare the performance of OTC stock trades across various price categories. The transaction value-weighted returns are presented in Table 4.8, Panel A and B. The results indicate that by a wide margin, stocks priced under $ 20 perform better than the rest on both a gross and net return basis. The lower priced stocks are more greatly impacted by commission costs which is a natural consequence of traders demanding a certain minimum commission fee for trading stocks.

The value weighted commission percentage decreases from 9.8 percent for the stocks priced at or under $1 to 0.75 percent for stocks valued at above $100. Another interesting finding is the low level of statistical significance associated with returns for stocks priced at or under $1, an indication of large dispersion of returns. This points out the large dispersion in observed round-trip returns on low priced OTC stocks. I hypothesize that since the average investor invests a minuscule portion of his securities in OTC securities, he/she does not mind accepting lottery-like outcomes in an attempt to capture a huge profit or upside unavailable in other stock markets. This could be an implication of behavioral portfolio theory developed by Shefrin and Statman (2000), wherein investors choose portfolios by considering expected wealth, desire for security and potential, aspiration levels, and probabilities of achieving aspiration levels. By investing a small portion of their portfolio in OTC securities, investors are willing to take a huge amount of risk for a shot at huge gains. Thus, I study the distribution of gross returns on OTC securities, presented in Table 4.9. I find that the medians are less than the means for all stocks priced at less than or equal to $100, indicating positively skewed returns. The measure of the skewness for each price category validates my conclusion. Interestingly, the maximum gross return earned on a round-trip trade for stocks valued under $1 was 11,775 percent. Similarly, high maximum returns are observed for all other round-trip trades, except for those on stocks valued at above $100. By any stretch of the imagination, this return is unattainable in stocks traded on the NYSE/NASDAQ/AMEX in a short period of time. However, the median shows that the majority of the trades do not perform very well. The median return for stocks priced at under $1 was -17.27 percent. In spite of huge commission costs associated with lower priced OTC stocks, I find that the net return distribution presented in Table 4.9, Panel B is similar to the gross return distribution. The results indicate the tremendous downside risk in OTC stocks priced under $100 but also reflect the probability of extremely high returns possible with very low probability.

## 10 Conclusion

Using Odean's (1998) discount brokerage data for the period January 1991 to November 1996, I find that trading in OTC stocks (i.e. stocks traded on Pinks Sheets or OTCBB) is a pervasive behavior with nearly half of the (32,108 out of 65,591) households holding a position in an OTC stock during the 71 month period. An examination of the roundtrip returns (from buy to sale of stock) shows that the returns for OTC stocks are positively skewed. There is a tremendous downside risk in OTC stocks priced under $100 but also reflect the probability of extremely high returns possible with very low probability. Perhaps this is one of the motivations for a large number of individuals investing in these stocks. The other could be a long term investment in a company that does not meet the listing standards or does not see benefits to being listed on the major stock exchanges (e.g. Nestlé). The average household's portfolio mix irrespective of their wealth tends to be fairly similar across a range of equity and fixed income instruments with the average investment in OTC securities being in the range of 2.16 to 3.15 percent.

Investing in OTC stocks is behavior commonly identified with less savvy and low income individuals taking a shot at the riches. Contrary to the characteristics associated with lottery buyers/gamblers, OTC investors are older, wealthier and more experienced at investing. These results are robust to alternate definition of OTC investors and statistical techniques used for handling missing data. My tests indicate that OTC-investor accounts are not play money accounts, which makes these conclusions more relevant.

Other than demographic characteristics, this study also examines the portfolio characteristics associated with OTC investors. I find that OTC investors tend to be more diversified relative to their peers. Thus, there is little evidence that OTC investors are generally risk seeking individuals concentrating their portfolios in a few stocks. They exhibit higher portfolio turnovers in their Non-OTC stocks in comparison to Non-OTC portfolios, an expected finding given the risk surrounding OTC investing individuals need to keep a close watch on market and trade whenever beneficial as opposed to taking a passive approach to investing. Also, investors who invest a larger portion of their portfolio in large cap stocks are less likely to buy OTC stocks. Micro capitalization stocks are fairly similar to OTC stocks. Expectedly, I find a positive relationship between the portion of portfolio allocated to micro caps and the propensity to own OTC stocks.

Overall, the results of this study provide evidence of pervasiveness of OTC investors. It also dispels popular notions of OTC investor being poor or less sophisticated. The results are largely consistent with Shefrin and Statman's (2000) behavioral portfolio theory, which posits that investors in general construct their portfolios as layered pyramids, where the bottom layers are designed for downside protection, while top layers are designed for upside potential.

## REFERENCES


Barber, Brad and Terrance Odean, 2001, "Boys will be boys: Gender, overconfidence, and common stock investment," *Quarterly Journal of Economics*, 116, 261–292.

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

Barber, Brad M. and Terrance Odean, 2006, All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *Review of Financial Studies*, Forthcoming.

Barber, Brad M., Terrance Odean, and Michal Strahilevitz, March 2004, Once Burned, Twice Shy: Naïve Learning, Counterfactuals and the Repurchase of Stocks Previously Sold, *Working Paper*.

Barberis, Nicholas, Andrei Schleifer, and Robert Vishny, 1998, A Model for Investor Sentiment, *Journal of Financial Economics*, 49, 307-343.

Barsky, Robert B, Juster, F. Thomas; Kimball, Miles S. and Shapiro, Matthew D. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics*, May 1997, 112 (2), 537-79.

Beatty, Randolph and Padma Kadiyala, 2003, "Impact of the Penny Stock Reform Act of 1990 on the Initial Public Offering Market," *Journal of Law and Economics*, 46, 517-542.

Benartzi, Shlomo and Richard Thaler, 2007, Heuristics and Biases in Retirement Savings Behavior, *Journal of Economic Perspectives*, 21(3), 81-104.

Benartzi, Shlomo, 2001, Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock, *Journal of Finance*, 56, 1747-1764.

Blume, Marshall E., and Irwin Friend, 1975, "The Asset Structure of Individual Portfolios and Some Implication for Utility Functions," *Journal of Finance*, 30, 585-603.

Blume, Marshall E., Jean Crockett, and Irwin Friend, 1974, "Stock Ownership in the United States: Characteristics and Trends," *Survey of Current Business*, 54(11), 16-40.

Boston Research Group, April 25, 2002, Enron Has Little Effect on 401(k) Participants' View of Company Stock, Boston: MA.

Bradley, Daniel J., John W. Cooney, Jr., Steven D. Dolvin, and Bradford D. Jordan, Spring 2006, "Penny Stock IPOs," *Financial Management*, 5-29.

Carhart, Mark M., 1997, On Persistence of Mutual Fund Performance, *Journal of Finance*, 52, 57-82.

Chan, Kalok, Vicentiu Covrig and Lilian Ng, 2005, What Determines the Domestic bias and Foreign Bias? Evidence from Mutual Fund Equity Allocations Worldwide, *Journal of Finance*, 60, 1495-1534.

Chen, Gongmeng, Kenneth A. Kim, John R. Nofsinger, and Oliver M. Rui, 2007, Trading Performance, Disposition Effect, Overconfidence, Representativeness Bias, and Experience of Emerging Market Investors, *Journal of Behavioral Decision Making*, 20(4), 425-451.

Chen, Gongmeng, Kenneth Kim, John R. Nofsinger and Oliver Rui, 2007, "Trading Performance, Disposition Effect, Overconfidence, Representativeness Bias, and Experience of Emerging Market Investors," *Journal of Behavioral Decision Making*, 20(4), 425-451.

Chen, Richard, 2003, "Chronology of Controversial Practices in a Microcap Company," *Institutional Investor*, 37, 60-68.

Clotfelter, Charles T. and Philip J. Cook, 1989, *Selling Hope: State Lotteries in America* (Harvard University Press, Cambridge, MA).

Cooper, Ian, and Evi Kaplanis, 1994, Home Bias in Equity Portfolios, Inflation Hedging, and International Capital Market Equilibrium, *Review of Financial Studies*, 7, 45-60.

Coval, Joushua D., and Tobias J. Moskowitz, 1999, Home Bias at Home: Local Equity Preference in Domestic Portfolios, *Journal of Finance*, 54, 2045-2073.

De Bondt, Werner F.M, 1993, Betting on Trends: Intuitive Forecasts of Financial Risk and Return, *International Journal of Forecasting*, 9, 355-371.

Deogun, Nikhil, 1997, "The Legacy: Robert Goizueta Led Coca-Cola Surge and Its Home Prospers," *Wall Street Journal,* October 20.

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake, 2007, Participant Reaction and the Performance of funds offered by 401(k) plans, *Journal of Financial Intermediation*, 16, 249-271.

Fama, Eugene F. and Kenneth R. French, 1993, Common Factor in the Returns of Stocks and Bonds, *Journal of Financial Economics*, 33, 3-56.

Feng, Lei and Mark S. Seasholes, 2005, "Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets?" *Review of Finance*, 9, 305-351.

French, Kenneth R., and James M. Poterba, 1991, Investor Diversification and International Equity Markets, *American Economic Review*, 81, 222-226.

Goetzmann, William N. and Alok Kumar, 2008, "Equity Portfolio Diversification," *Review of Finance*, 2008, 12, 433–463.

Goetzmann, William N., Lingfeng Li, and K. Geert Rouwenhorst, 2005, Long-term global market correlations, *Journal of Business*, 78, 1-38.

Griffin, Dale and Amos Tversky, 1992, The Weighting of Evidence and the Determinants of Confidence, *Cognitive Psychology*, 24, 411-435.

Grinblatt, Mark, and Matti Keloharju, 2000, The investment behavior and performance of various investor types: A study of Finland's unique data set, *Journal of Financial Economics, 55,* 43–67.

Grossman, Sanford J. and Joseph E. Stiglitz, 1980, On the Impossibility of Informationally Efficient Markets, *American Economic Review*, 70, 393-408.

Grubel, Herbert G., 1968, Internationally Diversified Portfolios, *American Economic Review*, 58, 1299-1314.

Hanke, Michael and Florian Hauser, 2008, "On the effects of stock spam emails," *Journal of Financial Markets*, 11, 57-83.

Hanson, Tim and Brian Richards, December 16, 2006, "Why we Love Wild Penny Stocks," *The Motley Fool, Fool.com*

Harris, Jeffrey H., Venkatesh Panchapagesan and Ingrid M. Werner, March 2008, "Off but not Gone: A Study of Nasdaq Delistings," *Working Paper*.

Heath, Chip, and Amos Tversky, 1991, Preference and Belief: Ambiguity and Competence in Choice under Uncertainty, *Journal of Risk and Uncertainty*, 4, 5-28.

Huberman, Gur, 2001, "Familiarity Breeds Investment," *Review of Financial Studies*, 14, 659-680.

Ivkovic, Zoran, and Scott Weisbenner, 2005, "Local Does as Local Is: Information Content of the Geography of Individual Investors' Common Stock Investments," *Journal of Finance*, 60, 267-306.

Jianakoplos, Nancy Ammon and Bernasek, Alexandra. "Are Women More Risk Averse?" *Economic Inquiry*, October 1998, 36(4), 620-30.

Kahneman, Daniel, and Amos Tversky, 1972, Subjective Probability: A Judgment of Representativeness, *Cognitive Psychology*, 3, 430-454.

Kennickell, Arthur B., Martha Starr-McCluer, and Annika E. Sunden, 1997, Family Finances in the US: Recent Evidence from the Survey of Consumer Finances, *Federal Reserve Bulletin*, 83, 1-24.

Kumar, Alok, 2009, "Who Gambles in the Stock Market?," *Journal of Finance*, Forthcoming.

Lakonishok, Josef, and Andre Shleifer, and Robert W. Vishny, 1994, Contrarian Investment, Extrapolation, and Risk, *Journal of Finance*, 69, 1541-1578.

Laura, Frieder, and Avanidhar Subrahmanyam, 2005, Brand Perceptions and the Market for Common Stock, *Journal of Financial and Quantitative Analysis*, 40, 57-85.

Levy, Haim, and Marshall Sarnat, 1970, International Diversification of Investment Portfolios, *American Economic Review*, 60, 668-675.

Lewis, Karen, 1999, Trying to Explain Home Bias in Equities and Consumption, *Journal of Economic Literature*, 37, 571-608.

Massa, Massimo, and Andrei Simonov, 2006, "Hedging, Familiarity and Portfolio Choice," *Review of Financial Studies*, 19, 633-685.

McLean, Bethany, May 15, 2000, "Making Little Cents," *Fortune*, 141, 446-447.

Merton, Robert C., 1987, A Simple Model of Capital Marker Equilibrium with Incomplete Information, *Journal of Finance*, 42, 483-510

Moggridge, Donald, 1983, *The Collected Writings of John Maynard Keynes: Volume XII, Economic Articles and Correspondence: Investment and Editorial*, Macmillan Cambridge University Press, London.

National Gambling Impact Study Commission. 1999. *National gambling impact study commission report*. (http://www.ngisc.gov)

Odean, Terrance, 1998, "Are Investors Reluctant to Realize Their Losses," *Journal of Finance*, 53, 1775-1798.

Odean, Terrance, 1998, Are Investors Reluctant to Realize Their Losses, *Journal of Finance*, 53, 1775-98.
Odean, Terrance, 1999, "Do Investors Trade Too Much?" *American Economics Review*, 89, 1279–1298.

Roane, Kit R., June 4, 2007, "Taking Penny Stocks out of the Boiler Room; New service aims to make 'pink sheet' issues a bit less risky" *U.S. News & World Report.*

Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York.

Schafer, J.L. (1997), *Analysis of Incomplete Multivariate Data*, New York: Chapman and Hall.

Scholes, Myron and Joseph Williams, 1977, "Estimating Betas for Nonsynchronous Data," *Journal of Financial Economics*, 5, 309-327.

Serrat, Angel, 2001, “A Dynamic Equilibrium Model of International Portfolio Holdings,” *Econometrica*, 69, 1467-1489.

Shefrin, Hersh and Meir Statman, 2000, “Behavioral Portfolio Theory,” Journal of Financial and Quantitative Analysis, 35(2), 127-150.

Shefrin, Hersh, 2005, *A Behavioral Approach to Asset Pricing*, Elsevier Academic Press.
Solnik, Bruno, 1974, An Equilibrium Model of the International Capital Market, *Journal of Economic Theory*, 8, 500-524.

Solt, Michael E., and Meir Statman, 1989, Good Companies, Bad Stocks, *Journal of Portfolio Management*, 15, 39-44.

Statman, Meir, 1987, How Many Stocks Make a Diversified Portfolio?, *Journal of Financial and Quantitative Analysis*, 22, 353-363.

Stulz, René, 1981, “On the Effect of Barriers to International Investment,” *Journal of Finance*, 36, 923-934.

Tesar, Linda, and Ingrid Werner, 1995, Home Bias and High Turnover, *Journal of International Money and Finance*, 14, 467-492.

Tversky, Amos, and Daniel Kahneman, 1974, Judgement under Uncertainity: Heuristics and Biases, *Science*, 185, 1124-1131.

Vissing-Jorgensen, Annette, 2003, Perspectives on Behavioral Finance: Does ‘Irrationality’ Disappear with Wealth? Evidence from Expectations and Actions, *NBER/Macroeconomics Annual,* 18(1), 139-94.

**Table 2.1**

**Descriptive Statistics on Price, Trade Size, and Commission**

The sample covers trading records of 62,942 households at a large discount brokerage house from January 1996 to November 1996. Commission is calculated as commission paid by the value of the trade.

| Variable | Mean | 25th Percentile | Median | 75th Percentile | Standard Deviation | No of Obs | No of households |
|---|---|---|---|---|---|---|---|
| Panel A : Purchase | | | | | | | |
| Price ($) | 31.00 | 11.00 | 23.00 | 40.00 | 115.77 | 1071182 | 55902 |
| Trade Size ($) | 11143 | 2513 | 4975 | 10500 | 31841 | 1071182 | 55902 |
| Commission (%) | 2.07 | 0.82 | 1.40 | 2.41 | 4.86 | 1071182 | 55902 |
| Panel B : Sale | | | | | | | |
| Price ($) | 31.13 | 12.00 | 24.00 | 41.00 | 110.15 | 876116 | 56997 |
| Trade Size ($) | 13630 | 2688 | 5725 | 13000 | 38087 | 876116 | 56997 |
| Commission (%) | 3.01 | 0.73 | 1.25 | 2.23 | 140.93 | 876116 | 56997 |
| Panel C : Securities not listed on NYSE/AMEX/NASDAQ | | | | | | | |
| Price ($) | 9.71 | 0.00 | 0.50 | 3.12 | 153.80 | 5307 | |

## Table 2.2

### Descriptive statistics of number, value, and average size of trades conducted by repurchasing households across size of stocks

The sample consists of the trading records of 22,971 households that repurchase a stock that was previously sold. These households are largely representative of the trading activity of the entire 62,942 households that traded in common stocks at a large discount brokerage house from Jan 1991 to November 1996. The *NYSE/AMEX/NASDAQ size deciles* were formed on 29 November, 2006 by extracting data on non-zero market capitalization stocks that were listed in on the CRSP database.* The market capitalization ranges for the various deciles is used to allocate traded stocks to different deciles. The day when a stock is last traded in the sample is used to determine the market capitalization of the traded stock. The *Not in CRSP* category refers to stocks that are traded on exchanges other than the NYSE/ AMEX/NASDAQ, for which no data is available in this study. *Mkt cap range* refers to the market capitalization range for each size decile. *Number of stocks traded* by households refers to stocks that might have been purchased or sold at the brokerage house. Panel C, shows repurchases, which are a subset of purchases presented in Panel A.

| | Not | NYSE/AMEX/NASDAQ decile cutoffs determined as on 29 Nov 1996 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | in CRSP | 1(Lowest) | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10(Highest) |
| Panel A: Market capitalization ranges and number of stock traded by households | | | | | | | | | | | |
| Mkt cap range ($ in millions) | N/A | 0-11 | 11-23 | 23-41 | 41-66 | 66-104 | 104-166 | 166-281 | 281-534 | 534-1438 | 1438-17173 |
| # of stocks traded by households | 4058 | 1402 | 993 | 990 | 826 | 894 | 909 | 862 | 930 | 930 | 944 |
| Panel B: Purchases | | | | | | | | | | | |
| % of Total Number of buy trades | 3.54 | 2.52 | 2.54 | 3.14 | 3.37 | 4.17 | 4.68 | 6.03 | 7.80 | 12.05 | 50.07 |
| % of Total dollar value of buys | 2.41 | 0.92 | 1.15 | 1.54 | 1.97 | 2.47 | 3.24 | 4.69 | 6.24 | 11.43 | 63.91 |
| Average Trade Size ($) | 8478 | 4539 | 5638 | 6094 | 7253 | 7372 | 8606 | 9674 | 9959 | 11804 | 15889 |
| Panel C: Sales | | | | | | | | | | | |
| % of Total Number of Sales | 3.41 | 2.06 | 2.26 | 2.78 | 3.06 | 3.84 | 4.45 | 5.91 | 7.71 | 12.14 | 52.26 |
| % of Total dollar value of Sales | 2.16 | 0.72 | 0.92 | 1.28 | 1.72 | 2.17 | 2.98 | 4.40 | 6.16 | 11.32 | 66.12 |
| Average Trade Size ($) | 9515 | 5280 | 6112 | 6923 | 8407 | 8480 | 10056 | 11170 | 11988 | 14000 | 18997 |
| Panel D: Repurchases (Subset of Purchases shown in Panel B) | | | | | | | | | | | |
| No of Stocks Repurchased | 405 | 299 | 392 | 465 | 496 | 511 | 538 | 577 | 612 | 670 | 812 |
| % of Total number of repurchases | 2.27 | 0.41 | 0.84 | 1.55 | 2.30 | 3.32 | 4.62 | 6.33 | 8.89 | 15.65 | 53.79 |
| % of Total dollar value of repurchases | 1.88 | 0.10 | 0.26 | 0.59 | 1.01 | 1.66 | 2.70 | 4.36 | 7.27 | 14.86 | 65.29 |
| Average Trade Size ($) | 16951 | 5061 | 6356 | 7816 | 8956 | 10211 | 11948 | 14097 | 16710 | 19405 | 24806 |

*There are 79 stocks traded by the repurchasing households, which have a zero market capitalization in the CRSP database. The descriptive statistics for trades in these stocks are not presented here as the trading activity in these stocks forms a negligible portion of total trades (less than .05% in value).

**Table 2.3**
**Descriptive statistics for repurchasing household categories based on frequency of repurchase**

The repurchase activity and characteristics of 22,971 households has been described based on frequency of repurchases. All statistics presented in this table are mean values for all households in a particular category. The mean monthly turnover for each household, calculated as per the methodology in Barber and Odean (2000), is the average of the purchase and sales turnover. *Monthly turnover* is the beginning-of-month market value of shares purchased in month t-1 (or sold in month t) divided by the total beginning-of month market value of shares held in month t.

| Category ( # of repurchase trades) | # of Households | Repurchases as %age of # of purchases | Repurchases as %age of total value of purchase | # of stocks Repurchased | # of stocks Traded | Mean Monthly Turnover (%) | Repurchased stocks as %age of all stocks traded |
|---|---|---|---|---|---|---|---|
| A ( 1 ) | 8543 | 17 | 20 | 1 | 12 | 7 | 19 |
| B ( 2 ) | 4027 | 18 | 21 | 2 | 17 | 9 | 19 |
| C ( 3 ) | 2312 | 19 | 23 | 2 | 20 | 10 | 20 |
| D ( 4 ) | 1505 | 19 | 24 | 3 | 23 | 11 | 21 |
| E ( 5 ) | 1032 | 19 | 24 | 4 | 26 | 12 | 21 |
| F ( 6-7 ) | 1383 | 21 | 26 | 5 | 29 | 14 | 23 |
| G ( 8-10 ) | 1221 | 22 | 28 | 6 | 34 | 15 | 24 |
| H ( 11-20 ) | 1597 | 24 | 30 | 9 | 44 | 18 | 27 |
| I ( 21-50 ) | 993 | 28 | 34 | 17 | 64 | 22 | 33 |
| J ( 51-100 ) | 261 | 32 | 38 | 32 | 95 | 27 | 40 |
| K ( 101-1181 ) | 97 | 36 | 42 | 70 | 199 | 35 | 43 |

## Table 2.4

## Analysis of time between sale and subsequent repurchase trade

This table shows the number of calendar days between sale and subsequent repurchase transaction, labeled as *Pre-repurchase period* in this table. Panel A provides a description among various repurchases categories formed on the basis of repurchase activity of these households. Panel B shows the percentage of repurchase trades that took place within various time intervals.

Panel A

| Category ( # of repurchase trades) | # of households | Pre-repurchase period Mean | Median | Standard Dev |
|---|---|---|---|---|
| A ( 1 ) | 8543 | 311 | 161 | 378 |
| B ( 2 ) | 4027 | 263 | 168 | 275 |
| C ( 3 ) | 2312 | 245 | 168 | 234 |
| D ( 4 ) | 1505 | 228 | 161 | 214 |
| E ( 5 ) | 1032 | 210 | 159 | 180 |
| F ( 6-7 ) | 1383 | 205 | 154 | 171 |
| G ( 8-10 ) | 1221 | 185 | 142 | 156 |
| H ( 11-20 ) | 1597 | 174 | 138 | 141 |
| I ( 21-50 ) | 993 | 147 | 125 | 100 |
| J ( 51-100 ) | 261 | 126 | 109 | 90 |
| K ( 101-1181 ) | 97 | 117 | 91 | 94 |

Panel B

| Pre-repurchase period (T) | %age of all repurchase trades |
|---|---|
| T <= 1 week | 17.21 |
| 1 week < T <= 1 month | 19.52 |
| 1 month < T <= 6 months | 35.04 |
| 6 months < T <= 1 Year | 12.58 |
| 1 Year < T <= 2 Years | 9.32 |
| 2 Years < T | 6.33 |

## Table 2.5

### Repurchase trades across various industry segments

This table shows the distribution of dollar values of repurchase trades across various industry segments and across the top three firms within each segment. Repurchases (in $ millions) shows the value of repurchases trades in each industry segment.* *% of Repurchase* refers to the percentage of dollar value of total repurchases, made in each industrial segment. *Firm No 1* refers to the firm in which highest dollar value of repurchases are made within a particular industry segment. Similarly the *Firm No 2* and *Firm No 3* account for the second and third highest portion of repurchases, respectively. The percentage provided in front of each firm shows the concentration of dollar value of repurchases in each firm within an industrial segment.

| *Industry Segment* | Repurchases ($ in millions) | % of Repurchase | Top 3 Firms (%age of total industry repurchase) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Firm No1 | | Firm No2 | | Firm No3 | |
| *Computer Hardware* | 477 | 16.44% | Micron Technology Inc | 17% | IBM | 14% | Cisco | 10% |
| *Electronic Equipment* | 296 | 10.18% | Intel | 29% | Texas Instruments Inc | 6% | LSI Logic Corp | 5% |
| *Computer Software* | 278 | 9.58% | Microsoft | 25% | Novell Inc | 9% | Netscape Comm. Corp | 6% |
| *Pharmaceutical Products* | 195 | 6.72% | Merck & Co | 16% | Amgen Inc | 12% | Glaxo Wellcome | 7% |
| *Retail* | 172 | 5.91% | Wal Mart | 21% | Home Depot | 12% | K Mart | 7% |
| *Trading (Finace)* | 115 | 3.94% | Charles Schwab Corp | 16% | Bank of America | 11% | Citicorp | 10% |
| *Communications* | 109 | 3.76% | Telefonos De Mexico SA | 28% | AT & T | 19% | MCI Communications | 8% |
| *Electrical Equipment* | 106 | 3.63% | U S Robotics Corp | 18% | Bay Networks Inc | 11% | Westinghouse Electric | 9% |
| *Automobiles and Trucks* | 92 | 3.18% | General Motors | 32% | Chrysler Corp | 29% | Ford Motor Co | 24% |
| *Business Services* | 85 | 2.92% | Compaq Computer Inc | 37% | Presstek Inc | 11% | Autodesk Inc | 6% |
| *Medical Equipment* | 77 | 2.65% | US Surgical Corp | 30% | Johnson & Johnson | 23% | Summit Technology Inc | 6% |
| *Petroleum and Natural Gas* | 66 | 2.28% | Exxon Corp | 14% | Occidental Petroleum | 11% | Texaco Inc | 6% |
| *Measuring and Control Equp.* | 66 | 2.27% | Iomega Corp | 43% | Hewlett Packard Corp | 37% | KLA Instruments Corp | 5% |
| *Transportation* | 49 | 1.70% | United Airlines | 20% | Southwest Airlines | 15% | US Air Group | 12% |
| *Restaurants, Hotels & Motels* | 49 | 1.68% | Callaway Golf Co | 12% | McDonalds Corp | 12% | Circus Circus Enterprises | 10% |
| *Wholesale* | 48 | 1.66% | Conner Peripherals Inc | 16% | Nike Inc | 15% | Snapple Beverage Corp | 7% |
| *Machinery* | 47 | 1.61% | Applied Materials Incs | 43% | IVAX Corp | 8% | Caterpillar Inc | 4% |
| *Banking* | 45 | 1.54% | Wells Fargo & Co | 11% | American Express Co | 10% | Chase Manhattan Corp | 9% |
| *Utilities* | 43 | 1.49% | PG & E | 7% | American Electric Power | 7% | Texas Utilities Co | 6% |
| *Recreation* | 43 | 1.49% | Motorola | 46% | EMC Corp | 26% | Zenith Electronics | 8% |
| *Consumer Goods* | 41 | 1.42% | General Electric Corp | 30% | Bristol Myers Squibb Co | 20% | Eastman Kodak Corp | 9% |
| *Healthcare* | 39 | 1.33% | U S Healthcare Inc | 17% | United Healthcare Corp | 7% | Novacare Inc | 6% |
| *Candy & Soda* | 37 | 1.28% | Pepsico Inc | 54% | Coca Cola Corp | 43% | Wrigley William Jr. Corp | 1% |
| *Precious Metals* | 34 | 1.17% | Homestake Mining Corp | 21% | Barrick Gold Corp | 19% | Placer Dome Inc | 11% |
| *Tobacco Products* | 34 | 1.16% | Phillip Morris Corp | 95% | U S T Inc | 3% | Loews Corp | 1% |
| *FoodProducts* | 28 | 0.96% | RJR Nabisco Hldg Corp | 24% | Archer Daniels Midland | 16% | Quaker Oats Corp | 8% |
| *Aircraft* | 26 | 0.88% | Boeing Corp | 87% | McDonnell Douglas Corp | 5% | United Technologies | 1% |
| *Other Industries* | 208 | 7.16% | | | | | | |
| | 2904 | 100.00% | | | | | | |

SIC codes for each stock were obtained from CRSP. The industry segment were formed according to SIC code industrial classification obtained from Kenneth R French's online data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). Stocks are classified into 49 industrial segments but only the top 27 segments have been reported as they account for 93% of repurchases.

**Table 2.6**

**Market risk of repurchased stocks**

This table provides the average estimate of beta for stocks traded within each familiarity quintile. The beta is estimated using the Scholes and Williams (1977) procedure over a time window of 250 trading days prior to each transaction. Thereafter the repurchase transaction value weighted beta is calculated for each repurchasing household. For each familiarity quintile the market risk (Beta) of repurchased stocks is determined by taking an equally weighted average of betas for all households within each familiarity quintile. The F-test tests for equality across all familiarity quintiles. The p-values for each statistic are presented in parenthesis.

| Familiarity Quintile | No of households | CAPM Beta |
|---|---|---|
| 1 (Lowest) | 4248 | 1.36 |
| 2 | 4563 | 1.39 |
| 3 | 4479 | 1.43 |
| 4 | 4459 | 1.46 |
| 5 (Highest) | 4477 | 1.47 |
| | | |
| 5-1 (High - Low) | | .11 (.00) |
| F Test | | 20.69 (.00) |

## Table 2.7

### Testing for Familiarity Vs Information hypotheses – Evidence from gross return performance of roundtrip repurchase transactions

This table shows the gross return for each roundtrip repurchase transactions (i.e., returns earned on the first sale transaction after repurchase). A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock. For each household repurchase value weighted gross return is computed. Thereafter the value weighted gross returns are averaged across all households assigned to a familiarity quintile to arrive at the figure presented below. The *Market Adjusted Abnormal Return* presented in Panel A is the difference between the return on the roundtrip trade and the return on the market ($R_i$-$R_m$). The *Industry Adjusted Return* is the difference between the return on the roundtrip trade and the value weighted return earned by all stocks in the same industry, excluding the repurchased stock. The 49 industry classifications are obtained from Kenneth R. French's online data library. Panel B presets abnormal returns based on 4 different asset pricing models. The *Jensen's alpha* in is computed using the CAPM $[(R_i-R_f)-\beta(R_m-R_f)]$. The beta is estimated using the Scholes and Williams (1977) procedure over a time window of 250 trading days prior to each transaction. The CRSP value weighted NYSE/AMEX/NASDAQ is used to proxy for market returns and the 1 month t-bill rate is used as the risk free rate*. The *Fama and French (1993) 3 factor abnormal return* is the difference between the observed and predicted return on a model that adds size and value factors to the CAPM $[(R_i-R_f)-\beta_1(R_m-R_f)-\beta_2(SMB)-\beta_3(HML)]$. The *Carhart (1997) 4 factor abnormal return* is the difference between the observed and predicted return on a model that adds a momentum factor to the Fama French 3 Factor Model $[(R_i-R_f)-\beta_1(R_m-R_f)-\beta_2(SMB)-\beta_3(HML)-\beta_4(WML)]$. Similarly the last model stated as *Carhart 4 factor plus Industry factor abnormal return* is the difference between the observed and predicted return on a model that adds an industry return factor to the Carhart Model $[(R_i-R_f)-\beta_1(R_m-R_f)-\beta_2(SMB)-\beta_3(HML)-\beta_4(WML)- \beta_5(IND)]$. The IND factor is the value weighted return earned by all stocks in the same industry, excluding the repurchased stock. The SMB, HML, WML and IND factors capture size, value, and momentum and industry effects respectively*. The coefficient in each of the above factor models except for the CAPM were estimated using OLS regressions over 250 trading days prior to each repurchase. The *number households* refers to the total number of households in each familiarity quintile, for which data is available for analysis. The *number of repurchases* refers to the total number of roundtrip repurchase trades made by the households within each familiarity quintile. *Calendar days* is the average number of days between the sale and purchase trade in a roundtrip transaction, for all households in a familiarity quintile. Panel C shows the average number of roundtrips made on a stock by households within various familiarity quintiles. The F-Test in Panel C tests for equality across all familiarity quintiles. The p-values for all statistics are presented in parenthesis.

| Panel A: Nominal, Market Adjusted and Industry Adjusted Returns | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Familiarity Quintile | # of households | # of repurchases | Calendar days | | Nominal Returns (%) | | Market Adjusted Returns (%) | | Industry Adjusted Returns (%) | |
| 1 (Lowest) | 2728 | 4684 | 318 | (.00) | 13.38 | (.00) | -0.61 | (.00) | -16.04 | (.00) |
| 2 | 3464 | 9630 | 290 | (.00) | 12.00 | (.00) | -0.71 | (.00) | -14.55 | (.00) |
| 3 | 3693 | 17663 | 250 | (.00) | 10.06 | (.00) | -0.75 | (.00) | -13.37 | (.00) |
| 4 | 3906 | 29381 | 213 | (.00) | 10.71 | (.00) | 1.54 | (.00) | -9.17 | (.00) |
| 5 (Highest) | 3878 | 51856 | 198 | (.00) | 9.27 | (.00) | 0.79 | (.00) | -10.38 | (.00) |
| F-Stat | | | 112.96 | (.00) | 7.07 | (.00) | 3.44 | -(.01) | 12.63 | (.00) |

| Panel B: Abnormal returns based on various factor models | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Familiarity Quintile | # of households | # of repurchases | Jensen's Alpha (%) | | Fama and French 3 Factor (%) | | Carhart 4 Factor 4 Factor (%) | | Carhart 4 factor plus Industy factor (%) | |
| 1 (Lowest) | 2679 | 4577 | -3.95 | (.00) | -3.36 | (.00) | -3.11 | (.00) | -11.67 | (.00) |
| 2 | 3410 | 9373 | -3.90 | (.00) | -3.42 | (.00) | -3.65 | (.00) | -11.87 | (.00) |
| 3 | 3651 | 17172 | -3.19 | (.00) | -2.44 | (.00) | -2.77 | (.00) | -10.25 | (.00) |
| 4 | 3867 | 28360 | -0.84 | (.12) | -0.39 | (.48) | -0.23 | (.69) | -7.14 | (.00) |
| 5 (Highest) | 3852 | 50198 | -1.18 | (.01) | -0.52 | (.25) | -0.26 | (.59) | -7.01 | (.00) |
| F-Stat | | | 6.17 | (.00) | 6.11 | (.00) | 6.90 | (.00) | 10.94 | (.00) |

| Panel C: Roundtrips by households in every stock held | | |
|---|---|---|
| Familiarity Quintiles | # of roundtrips in each stock | |
| 1 (Lowest) | 1.05 | (.00) |
| 2 | 1.12 | (.00) |
| 3 | 1.20 | (.00) |
| 4 | 1.36 | (.00) |
| 5 (Highest) | 1.81 | (.00) |
| | | |
| 5-1 (High - Low) | 0.76 | (.00) |
| F Test | 1710.37 | (.00) |

*The 1 month T-bill rate (from Ibbotson Associates) and all other factor except IND have been obtained from Kenneth R. French’s website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research

## Table 2.8

## Testing for Representativeness amongst Repurchasing Households: Evidence from Roundtrip Trade performance of Non Repurchased Stocks Vs Repurchased Stocks

The sample for this table includes all the roundtrip trades made by repurchasing households. A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock. “Non Repurchased Stock” refers to stocks that are not repurchased after the first roundtrip. “Repurchased Stock” refers to stocks that are repurchased after the first roundtrip. For “Non Repurchased Stock” there is only one round-trip possible and this is used for the purpose of analysis. For “Repurchased Stock” the roundtrip trade prior to every repurchase in a particular stock is considered for analysis. The column “difference” refers to the difference between the repurchased and non repurchased stock at the households level. For each variable for every repurchasing household a repurchase value weighted computation is made. Thereafter the value weighted number is averaged across all households assigned to a familiarity quintile to arrive at the value of the variable presented in the tables below. The *average number of roundtrips* is the average number of roundtrip transactions for all households in a particular familiarity quintile. The *average calendar days per round trip* is the average number of days between the sale and purchase trade in a roundtrip transaction, for all households in a familiarity quintile. Please refer to Table 7 or the text for a description of performance measures. The p-values for all statistics are presented in parenthesis.

Panel A: Summary and Gross Returns

| | | # of roundtrips | | Roundtrip Calendar days | | | Gross Returns (%) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Familiarity Quintile | # of Households | Non-Repurchase Stock (1) | Repurchased Stock (2) | Non-Repurchase Stock (1) | Repurchased Stock (2) | Difference (1) - (2) | Non Repurchased Stock (1) | | Repurchased Stock (2) | | Difference (1) - (2) | |
| 1 (Lowest) | 3510 | 1.84 | 17.53 | 349 | 238 | 110 | 28.01 | (.00) | 36.86 | (.00) | -8.87 | (.00) |
| 2 | 3802 | 3.07 | 13.64 | 321 | 214 | 107 | 23.36 | (.00) | 32.71 | (.00) | -9.33 | (.00) |
| 3 | 3736 | 5.30 | 13.61 | 275 | 189 | 86 | 20.02 | (.00) | 27.19 | (.00) | -7.16 | (.00) |
| 4 | 3683 | 8.41 | 12.18 | 233 | 159 | 74 | 14.94 | (.00) | 24.05 | (.00) | -9.10 | (.00) |
| 5 (Highest) | 2795 | 18.18 | 10.01 | 186 | 127 | 60 | 9.00 | (.00) | 18.31 | (.00) | -9.32 | (.00) |

Panel B: Market Adjusted and Industry Adjusted Gross Returns (%)

| Familiarity Quintile | Market Adjusted Gross Returns (%): Non Repurchased Stock (1) | | Repurchased Stock (2) | | Difference (1) - (2) | | Industry Adjusted Gross Returns (%): Non Repurchased Stock (1) | | Repurchased Stock (2) | | Difference (1) - (2) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 (Lowest) | 12.51 | (.00) | 26.54 | (.00) | -14.05 | (.00) | -4.15 | (.00) | 13.56 | (.00) | -17.76 | (.00) |
| 2 | 9.17 | (.00) | 23.35 | (.00) | -14.15 | (.00) | -5.88 | (.00) | 11.29 | (.00) | -17.16 | (.00) |
| 3 | 7.94 | (.00) | 18.93 | (.00) | -10.97 | (.00) | -5.80 | (.00) | 7.66 | (.00) | -13.40 | (.00) |
| 4 | 4.70 | (.00) | 16.90 | (.00) | -12.20 | (.00) | -6.62 | (.00) | 7.72 | (.00) | -14.32 | (.00) |
| 5 (Highest) | 0.81 | (.08) | 12.55 | (.00) | -11.76 | (.00) | -7.47 | (.00) | 4.79 | (.00) | -12.48 | (.00) |

Panel C: Abnormal Gross Returns based on various factor models (%)

| Familiarity Quintile | Jensen's Alpha Difference (%) (1) - (2) | | Fama French 3 Factor Difference (%) (1) - (2) | | Carhart 4 Factor Difference (%) (1) - (2) | | Carhart 4 factor plus Industry Factor Difference (%) (1) - (2) | |
|---|---|---|---|---|---|---|---|---|
| 1 (Lowest) | -13.59 | (.00) | -12.96 | (.00) | -13.73 | (.00) | -15.07 | (.00) |
| 2 | -13.66 | (.00) | -13.59 | (.00) | -14.16 | (.00) | -14.43 | (.00) |
| 3 | -10.73 | (.00) | -10.56 | (.00) | -11.24 | (.00) | -10.87 | (.00) |
| 4 | -11.25 | (.00) | -11.24 | (.00) | -11.80 | (.00) | -11.61 | (.00) |
| 5 (Highest) | -11.49 | (.00) | -11.29 | (.00) | -11.47 | (.00) | -10.98 | (.00) |

## Table 2.9

### Logistic regressions modeling the decision to repurchase a stock for households conducting at least one repurchase

The table below presents the coefficient estimates for various logistic regressions that model the decision to repurchase a stock for households conducting at least one repurchase. The data set used for these models consists of roundtrip trades made on repurchased stocks and non-repurchased stocks used in section 8.1. Only those repurchase that take place within 1 calendar (ie 250 trading days) year from the last sale of a stock by a household are utilized in this analysis. The dependent variable is a dichotomous variable takes a value 1 if a repurchase trade is conducted on a stock and a value 0 if a stock is never repurchased. Model 1 has only stock characteristics as independent variables. Model 2 has only investor characteristics as independent variables. Model 3 has both stock and investor characteristics as independent variables. *Positive Net Market Adjusted Returns* is a dummy variable that takes the value 1 for stock that earned positive market adjusted returns in its last roundtrip repurchase and 0 otherwise. *Technology Sector* is a dummy variable that takes the value 1 for a stock lies in the technology sector and 0 otherwise. *Large Stock* is a dummy variable that takes the value 1 if the stock lies in the top market capitalization decile of all NYSE/AMEX/NASDAQ stocks and 0 otherwise. *Extreme Positive Previous Day Return* is a dummy variable that takes the value 1 if the stock returns lies in the highest previous day return decile for all stocks listed on the NYSE/AMEX/NASDAQ and 0 otherwise. The discount brokerage house classifies its clients into three categories: (1) *Affluent Traders*: Households with more that $100,000 in equities at any point of time (2) *Active Traders*: Households that make more than 48 trades in any year (3) *General Traders*: Households that not classified as affluent or active traders. Client segments are dummy variables that are coded treating the General traders as the reference category. The data on their knowledge of investments and net worth is self reported by the investors. The *age* dummy is obtained from a demographic database. *Knowledge* is a dummy variable that takes the value 1 if the investor considers himself/herself to have an extensive or good knowledge of investments and 0 if the investor considers himself/herself to have limited or no knowledge. *Size to Net worth ratio* is the ratio of the average monthly investment in common stock divided by the self reported net worth of the client. To account for any time specific clustering of repurchases over the 71 month period, 70 dummy variables are used to, the estimates for which are not reported to conserve space. The p-values for all statistics are presented in parenthesis.

| | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|
| Independent Variables | Estimates | Odds Ratio | Estimates | Odds Ratio | Estimates | Odds Ratio |
| Intercept | -1.9260 (.00) | | -1.4589 (.00) | | -2.5809 (.00) | |
| *Stock Characteristics* | | | | | | |
| Positive Net Market Adjusted Returns | 0.8213 (.00) | 2.273 | | | 0.8872 (.00) | 2.428 |
| Technology Sector | 0.6084 (.00) | 1.838 | | | 0.6421 (.00) | 1.900 |
| Large Stock | 0.4807 (.00) | 1.617 | | | 0.5511 (.00) | 1.735 |
| Extreme Positive Previous Day Return | 0.2004 (.00) | 1.222 | | | 0.1690 (.00) | 1.184 |
| *Investor Characteristics* | | | | | | |
| Client segment: Affluent Vs General Trader | | | -0.1813 (.00) | 0.834 | -0.2065 (.00) | 0.813 |
| Active Vs General Trader | | | 0.8242 (.00) | 2.280 | 0.8460 (.00) | 2.330 |
| Knowledge: High Vs Low | | | -0.0292 (.04) | 0.971 | -0.0023 (.88) | 0.998 |
| Age | | | -0.0023 (.00) | 0.998 | -0.0013 (.00) | 0.999 |
| Size to Net Worth Ratio | | | 0.0021 (.00) | 1.002 | 0.0023 (.00) | 1.002 |
| # of Independent Variables * | 74 | | 75 | | 79 | |
| # of observations: | | | | | | |
| Repurchases (Dependant variable=1) | 122,206 | | 35,208 | | 35,206 | |
| Non-Repurchases (Dependant variable=0) | 330,180 | | 88,386 | | 88,298 | |
| Total | 452,386 | | 123,594 | | 123,504 | |
| Wald Statistic (Chi-square p values) | 27,688 (.00) | | 5,265 (.00) | | 12,090 (.00) | |
| Psuedo R-squared | 0.0936 | | 0.0644 | | 0.1557 | |

* Including 70 dummy variables for the 71 month period between January, 1991 and November 1996

## Table 2.10

## Repurchases: Robustness Checks

This table provides evidence to test if repurchases are essentially gambles made by investors and if the trading accounts analyzed in this study are relevant to an investor's wealth. Panel A presents the percentage of average value of repurchases and average monthly investment in common stocks invested in Retirement (IRAs and Keoghs) and Non-Retirement accounts for households that hold both accounts at the discount brokerage house. Households that have both Retirement and Non-Retirement accounts constitute roughly half of households that made repurchases. Panel B presents the distribution of households among various Size to Net-Worth (SNR) deciles with each familiarity quintile. The *SNR ratio* is average monthly position in common stock divided by self reported net worth of each household. The *SNR quintiles breakpoints* are determined by measuring the SNR ratios for all the households that traded in common stocks irrespective of whether they did or did not make repurchases.

Panel A: Repurchasing households with both Retirement and Non-Retirement accounts

| Familiarity | # of | Repurchases | | Monthly Common Stock Position | |
|---|---|---|---|---|---|
| Quintile | Households | Retirement | Non-Retirement | Retirement | Non-Retirement |
| 1 | 2538 | 41.87% | 58.13% | 41.61% | 58.39% |
| 2 | 2487 | 42.01% | 57.99% | 41.47% | 58.49% |
| 3 | 2258 | 41.87% | 58.13% | 46.28% | 53.72% |
| 4 | 2205 | 41.64% | 58.36% | 42.60% | 57.40% |
| 5 | 1875 | 38.27% | 61.73% | 40.16% | 59.84% |

Panel B: Intersection of Size to Net Worth Ratio (SNR) and Familiarity quintiles

| | | | SNR Quintiles formed with all households that traded common stocks | | | | |
|---|---|---|---|---|---|---|---|
| Familiarity | # of | Mean | 1 | 2 | 3 | 4 | 5 |
| Quintiles | Households | SNR | SNR <= .21 | .21< SNR <= .48 | .48 < SNR <= .87 | .87 < SNR <= 1.67 | SNR > 1.67 |
| 1 (Lowest) | 2197 | 2.48 | 12.72% | 14.24% | 17.89% | 23.30% | 31.85% |
| 2 | 2138 | 1.74 | 13.80% | 17.01% | 19.88% | 24.18% | 25.12% |
| 3 | 1934 | 1.99 | 16.75% | 17.78% | 20.33% | 21.73% | 23.41% |
| 4 | 1842 | 1.66 | 19.43% | 19.37% | 18.86% | 21.20% | 21.14% |
| 5 (Highest) | 1598 | 1.34 | 22.35% | 19.31% | 19.64% | 19.11% | 19.58% |

## Table 2.11

### Buy and hold returns versus realized returns during the period between the first sale and the last repurchase

This table presents buy and hold returns, gross returns and net returns earned during the period between the first time a stock is sold and the last time it is repurchased by a household. For intermediate periods between a sale and subsequent repurchase of a stock, the gross and net returns are substituted with the prevailing risk free rate during that period. Thus, stocks that are repurchased only once have returns equivalent to the risk free rate. Only those repurchases that took place with 1 year from the last sale have been utilized in this analysis. The returns in this table are calculated in three steps. First, the returns are calculated for each stock repurchased by a household. Second, the average value of all repurchase transactions made by a household in each stock is used to compute a value weighted return for all stocks repurchased by each household. Finally, value weighted returns are averaged across all households assigned to a familiarity quintile to arrive at the figure presented below. The difference columns represent the difference between the buy and hold returns and realized returns (gross or net). The p-values for all statistics are presented in parenthesis.

| Familiarity Quintiles | # of households | Time | Buy and Hold Returns (1) | | Gross Returns (2) | | Difference (1) - (2) | | Net Returns (3) | | Difference (1) - (3) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 (Lowest) | 3397 | 117 | 7.25 | (.00) | 5.83 | (.00) | 1.42 | (.07) | 5.38 | (.00) | 1.86 | (.02) |
| 2 | 3639 | 108 | 9.76 | (.00) | 10.57 | (.00) | -0.82 | (.35) | 8.81 | (.00) | 0.94 | (.44) |
| 3 | 3815 | 103 | 15.77 | (.00) | 16.42 | (.00) | -0.66 | (.81) | 13.91 | (.00) | 1.86 | (.43) |
| 4 | 3643 | 95 | 15.52 | (.00) | 18.82 | (.00) | -3.31 | (.01) | 15.38 | (.00) | 0.14 | (.89) |
| 5 (Highest) | 4026 | 94 | 24.99 | (.00) | 30.25 | (.00) | -5.28 | (.01) | 24.61 | (.00) | 0.36 | (.85) |

## Table 2.12

### Net Return performance of Repurchase trades

This table shows the net return for each roundtrip repurchase transactions (ie. returns earned on the first sale transaction after repurchase). The net returns are obtained after adjusting gross returns for commission costs. A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock. For each household repurchase value weighted net return is computed. Thereafter the value weighted net returns are averaged across all households assigned to a familiarity quintile to arrive at the figure presented below. Please refer to Table 7 or the text for a description of performance measures. The *number households* refers to the total number of households in each familiarity quintile, for which data is available for analysis. The *number of repurchases* refers to the total number of roundtrip repurchase trades made by the households within each familiarity quintile. *Calendar days* is the average number of days between the sale and purchase trade in a roundtrip transaction, for all households in a familiarity quintile. Panel C shows the average number of roundtrips made on a stock by households within various familiarity quintiles. The F-Test in Panel C tests for equality across all familiarity quintiles. The p-values for all statistics are presented in parenthesis.

| Panel A: Nominal, Market Adjusted and Industry Adjusted Returns | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Familiarity Quintile | No of households | No of repurchases | Calendar days | | Nominal Returns (%) | | Market Adjusted Returns (%) | | Industry Adjusted Returns (%) | |
| 1 (Lowest) | 2728 | 4684 | 318 | (.00) | 9.81 | (.00) | -4.19 | (.00) | -19.61 | (.00) |
| 2 | 3464 | 9630 | 290 | (.00) | 8.70 | (.00) | -4.01 | (.00) | -17.84 | (.00) |
| 3 | 3693 | 17663 | 250 | (.00) | 7.06 | (.00) | -3.75 | (.00) | -16.37 | (.00) |
| 4 | 3906 | 29381 | 213 | (.00) | 7.96 | (.00) | -1.21 | (.02) | -11.92 | (.00) |
| 5 (Highest) | 3878 | 51856 | 198 | (.00) | 6.87 | (.00) | -1.61 | (.00) | -12.78 | (.00) |
| **Panel B: Abnormal returns based on various factor models** | | | | | | | | | | |
| Familiarity Quintile | No of households | No of repurchases | Jensen's Alpha (%) | | Fama and French 3 Factor (%) | | Carhart 4 Factor 4 Factor (%) | | Carhart 4 factor plus Industy factor (%) | |
| 1 (Lowest) | 2679 | 4577 | -7.54 | (.00) | -6.96 | (.00) | -6.71 | (.00) | -15.27 | (.00) |
| 2 | 3410 | 9373 | -7.20 | (.00) | -6.72 | (.00) | -6.94 | (.00) | -15.17 | (.00) |
| 3 | 3651 | 17172 | -6.22 | (.00) | -5.47 | (.00) | -5.80 | (.00) | -13.28 | (.00) |
| 4 | 3867 | 28360 | -3.60 | (.00) | -3.15 | (.00) | -2.99 | (.00) | -9.90 | (.00) |
| 5 (Highest) | 3852 | 50198 | -3.58 | (.00) | -2.91 | (.00) | -2.65 | (.00) | -9.40 | (.00) |

*The 1 month T-bill rate (from Ibbotson Associates) and all other factor except IND have been obtained from Kenneth R. French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research

## Table 2.13

## Impact of Familiarity on Diversification

This table shows estimates of various diversification measures across various familiarity quintiles. The STK measure is the number of stock in a portfolio. The WSQ measure is the sum of the square of the proportion invested in each security. The IND measure is the number of industries across which the household diversifies.39 The IWSQ measure is the sum of the square of the proportion invested in each industry. The NV measure is calculated as follows:

$$NV_{ewp} = \frac{\sigma_p^2}{\overline{\sigma}^2} = \frac{1}{N} + \left(\frac{N-1}{N}\right)\left(\frac{\overline{\text{cov}}}{\overline{\sigma}^2}\right)$$

where $\overline{\text{cov}}$ is the average covariance between stocks in the portfolio, $\sigma_p^2$ is the variance of the portfolio, and $\overline{\sigma}^2$ is the average variance of all stocks in the portfolios. The NVN measure is equal to the first term in the NV measure (ie. 1/N). The average correlation measure presents the average correlation for all portfolio that have at least 2 stocks. For each of the above measures a monthly average for each household is calculated. Thereafter these measures are averaged across all households within various familiarity quintiles to obtain the estimates presented in this table. The F Test tests for the equality of mean across familiarity quintiles. The p-values for all statistics are presented in parenthesis.

Panel A

| Familiarity Quintiles | Number of Stocks (STK) | | Weighted Square (WSQ) | | Number of Industries (IND) | | Industry Weighted Square (IWSQ) | |
|---|---|---|---|---|---|---|---|---|
| 1 (Lowest) | 8.14 | (.00) | 0.39 | (.00) | 3.87 | (.00) | 0.51 | (.00) |
| 2 | 5.50 | (.00) | 0.49 | (.00) | 3.11 | (.00) | 0.60 | (.00) |
| 3 | 4.69 | (.00) | 0.55 | (.00) | 2.78 | (.00) | 0.65 | (.00) |
| 4 | 3.92 | (.00) | 0.60 | (.00) | 2.45 | (.00) | 0.70 | (.00) |
| 5 (Highest) | 2.83 | (.00) | 0.72 | (.00) | 1.93 | (.00) | 0.80 | (.00) |
| 5-1 (High - Low) | -5.31 | (.00) | 0.34 | (.00) | -1.94 | (.00) | 0.29 | (.00) |
| F Test | 488.38 | (.00) | 1120.79 | (.00) | 943.53 | (.00) | 1044.60 | (.00) |

Panel B

| Household Category | Normalized Variance (NV) | | NV Naïve (NVN) | | Average Correlation (Portfolios >1 stock) | |
|---|---|---|---|---|---|---|
| Non-Repurchase | 0.71 | (.00) | 0.62 | (.00) | 0.2688 | (.00) |
| Familiarity Quintiles | | | | | | |
| 1 (Lowest) | 0.46 | (.00) | 0.32 | (.00) | 0.2542 | (.00) |
| 2 | 0.54 | (.00) | 0.41 | (.00) | 0.2551 | (.00) |
| 3 | 0.59 | (.00) | 0.47 | (.00) | 0.2569 | (.00) |
| 4 | 0.63 | (.00) | 0.53 | (.00) | 0.2595 | (.00) |
| 5 (Highest) | 0.74 | (.00) | 0.66 | (.00) | 0.2778 | (.00) |
| 5-1 (High - Low) | 0.2780 | (.00) | 0.34 | (.00) | 0.0235 | (.00) |
| F Test | 1013.00 | (.00) | 1010.82 | (.00) | 23.08 | (.00) |

[39] There are total number of 10 industry sectors that have been classified according to the definitions available Kenneth R. French's online data library.

## Table 2.14

## Impact of Investor Characteristics on Observed Diversification across Familiarity Quintiles

This table presents statistics for various investor groups formed on the basis of familiarity quintiles and investor categories assigned by the brokerage house. The discount brokerage house at which the investors maintain accounts categorizes households into three categories: (1) Affluent Trades: Households with more that $100,000 in equities at any point of time (2) Active Traders: Households that make more than 48 trades in any year (3) General Traders: Households that are neither classified as affluent or active traders. Panel A presents the percentage of Active, Affluent and General Investors in each household category. Panel B presents the average number of stocks held by households in various investor groups. Panel C presents the average number of industries across which households diversify their portfolios. The F Test tests for the equality of mean across familiarity quintiles. The p-values for all statistics are presented in parenthesis.

| Panel A | | | | |
|---|---|---|---|---|
| Household Category | No of Households | % age of Investor Composition | | |
| | | Active | Affluent | General |
| Non-Repurchase | 39973 | 2% | 17% | 81% |
| Familiarity Quintiles | | | | |
| 1 (Lowest) | 4594 | 27% | 23% | 50% |
| 2 | 4596 | 21% | 20% | 59% |
| 3 | 4597 | 20% | 17% | 63% |
| 4 | 4598 | 22% | 12% | 66% |
| 5 (Highest) | 4599 | 18% | 11% | 70% |

Panel B

| Household Category | Number of Stocks (STK) | | |
|---|---|---|---|
| | Active | Affluent | General |
| Non-Repurchase | 5.34 | 5.06 | 2.63 |
| Familiarity Quintiles | | | |
| 1 (Lowest) | 11.69 | 8.78 | 5.96 |
| 2 | 8.60 | 6.47 | 4.09 |
| 3 | 7.77 | 5.63 | 3.43 |
| 4 | 6.56 | 4.48 | 2.93 |
| 5 (Highest) | 5.29 | 3.18 | 2.13 |

Panel C

| Household Category | Number of Industries (IND) | | |
|---|---|---|---|
| | Active | Affluent | General |
| Non-Repurchase | 2.88 | 2.87 | 1.94 |
| Familiarity Quintiles | | | |
| 1 (Lowest) | 4.45 | 4.19 | 3.41 |
| 2 | 3.88 | 3.56 | 2.70 |
| 3 | 3.67 | 3.20 | 2.37 |
| 4 | 3.30 | 2.75 | 2.12 |
| 5 (Highest) | 2.83 | 2.18 | 1.66 |

## Table 3.1

### Observed portfolio allocation Vs Expected Portfolio Allocation for households in each state

This table presents the observed versus expected allocation (both in terms of number and value of stocks) to in-state direct utility companies within the direct utility portfolios held by households residing in various states. Each household's observed value (or number) allocation is computed by averaging the percentage value (or number) of the total value (or number of stocks) of its direct utility portfolio invested in local (in-state direct utility stocks over the number of months in which it held direct utility stocks. The observed allocation (A) for each state is computed by averaging the portfolio allocations for all stocks residing in the state. The expected value (B) allocation for each state is equal to the market capitalization of all companies operating within the state divided by the total market capitalization of all direct utility stocks. Similarly, the expected number allocation for each state is equal to the number of all companies operating within the state divided by the number of direct utility stocks in the United States. Finally, according to the US Census 1990 classification, I aggregated the observed and expected allocation for households across the following four regions: Midwest, Northeast, West and South. The two hypotheses (i.e. A=B and A=4B) are tested using paired right sided t-tests.

| Region | # of households | Household Allocation In-state (A) | Expected Allocation In-state (B) | *t statistic* (H0: A=B) | | *t statistic* (H0: A= 4B) | |
|---|---|---|---|---|---|---|---|
| Panel A: Allocation of $ amounts | | | | | | | |
| Midwest | 1673 | 18.93% | 2.34% | 20.61 | *** | 11.97 | *** |
| Northeast | 1815 | 21.22% | 2.76% | 21.98 | *** | 12.15 | *** |
| South | 2476 | 14.77% | 3.19% | 20.44 | *** | 3.58 | *** |
| West | 4329 | 20.70% | 4.26% | 32.71 | *** | 7.18 | *** |
| Panel B: Allocation among # of firms | | | | | | | |
| Midwest | 1673 | 18.59% | 2.74% | 19.94 | *** | 9.70 | *** |
| Northeast | 1815 | 20.83% | 3.46% | 20.92 | *** | 8.49 | *** |
| South | 2476 | 14.52% | 2.16% | 22.09 | *** | 10.55 | *** |
| West | 4329 | 20.39% | 3.43% | 34.90 | *** | 13.64 | *** |

* Significance at 10% level
** Significance at 5% level
*** Significance at 1 % level

**Table 3.2**

**Summary of Multiple Logistic Regression results modeling the decision the purchase direct utility stocks**

This table presents the average estimates and other summary results for 170 (i.e. the number of utility firms purchased by any household during the 71 month period) separate logistic regressions conducted to model the decision to purchase each direct utility firm. Each regression has 7,149 observations, which represent the number of households that purchase any utility stock during the 71 month period. Only the first purchase made by a household in each stock is considered in each regression to remove possible impacts of clustered observations. Each Model has results for two sub-parts. Part (A) presents the estimates for all regressions that converged during maximum likelihood estimation. Part (B) presents the estimates for all regressions that converged during maximum likelihood estimation and replaces the coefficients that were insignificant at $\alpha=0.05$ with zeros. Model 1 models the logit of a decision to purchase a stock x on a Local (In-state) Dummy. The Local (In-state) Dummy acquires a value of 1 if the household resides in the same state as the operational location of the company and 0 otherwise. Model 2 models the logit of a decision to purchase a stock x on a Local (In-state) Dummy and Previous 1 month Industry Adjusted Return. The Previous 1 month Industry Adjusted Return is used to control for possible impacts of investor extrapolating past returns while deciding to purchase a utility stock. For household that did not purchase a particular utility stock the previous 1 month Industry Adj Return is calculated using the first date when the household purchased any utility stock as the reference date. Model 3 extends Model 2 with an affluent dummy to capture wealth affects. The Affluent Dummy takes a value 1 if the household is deemed affluent by the brokerage house and 0 otherwise. Model 4 & 5 replaces the affluent dummy with diversification measures. The P-values for all tests are presented in parentheses. P-values for means are based on a one-sided (right side) t-test. P-values for medians are based on the sign rank test under which the null hypothesis is that median equals 0.

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A: Summary of Regressions** | | | | | | | | | | |
| # of regressions (firms) converged | 154 | | 154 | | 145 | | 153 | | 154 | |
| In-state Residents across regressions: | | | | | | | | | | |
| Mean | 461 | | 461 | | 471 | | 464 | | 461 | |
| Median | 228 | | 228 | | 233 | | 228 | | 228 | |
| **Panel B: Mean coefficients for coefficient estimates** | | | | | | | | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| | A | B | A | B | A | B | A | B | A | B |
| Intercept | -5.17 (.00) | -5.17 (.00) | -5.20 (.00) | -5.20 (.00) | -5.12 (.00) | -5.12 (.00) | -5.35 (.00) | -5.17 (.00) | -4.93 (.00) | -4.93 (.00) |
| Local (In-state) Dummy | 1.38 (.00) | 1.54 (.00) | 1.37 (.00) | 1.53 (.00) | 1.34 (.00) | 1.44 (.00) | 1.41 (.00) | 1.46 (.00) | 1.37 (.00) | 1.53 (.00) |
| Prior 1 month Ind Adj Return | | | -0.04 (.00) | -0.03 (.00) | -0.04 (.00) | -0.03 (.00) | -0.05 (.00) | -0.03 (.00) | -0.04 (.00) | -0.03 (.00) |
| Affluent Investor Dummy | | | | | 0.14 (.00) | 0.05 (.02) | | | | |
| # of stocks (Diversification measure) | | | | | | | 0.02 (.00) | 0.01 (.00) | | |
| Avg Correlation (Diversification measure) | | | | | | | | | -0.78 (.00) | -0.29 (.00) |
| **Panel C: Median coefficients for coefficient estimates** | | | | | | | | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| | A | B | A | B | A | B | A | B | A | B |
| Intercept | -5.07 (.00) | -5.07 (.00) | -5.05 (.00) | -5.05 (.00) | -5.05 (.00) | -5.05 (.00) | -5.14 (.00) | -5.08 (.00) | -4.93 (.00) | -4.93 (.00) |
| Local (In-state) Dummy | 1.62 (.00) | 1.50 (.00) | 1.62 (.00) | 1.46 (.00) | 1.59 (.00) | 1.41 (.00) | 1.64 (.00) | 1.43 (.00) | 1.62 (.00) | 1.47 (.00) |
| Prior 1 month Ind Adj Return | | | -0.05 (.00) | 0.00 (.00) | -0.05 (.00) | 0.00 (.00) | -0.05 (.00) | 0.00 (.00) | -0.05 (.00) | 0.00 (.00) |
| Affluent Investor Dummy | | | | | 0.16 (.00) | 0.00 (.03) | | | | |
| # of stocks (Diversification measure) | | | | | | | 0.01 (.00) | 0.01 (.00) | | |
| Avg Correlation (Diversification measure) | | | | | | | | | -0.50 (.00) | 0.00 (.00) |
| **Panel D: Proportion of positive or negative estimates (single sided p-values based on binomial test)** | | | | | | | | | | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| | A | B | A | B | A | B | A | B | A | B |
| Local (In-state) Dummy | 0.93 (.00) | 0.72 (.00) | 0.92 (.00) | 0.71 (.00) | 0.93 (.00) | 0.71 (.00) | 0.93 (.00) | 0.70 (.00) | 0.92 (.00) | 0.71 (.00) |
| Prior 1 month Ind Adj Return | | | 0.77 (.00) | 0.32 (.00) | 0.77 (.00) | 0.34 (.00) | 0.77 (.00) | 0.33 (.00) | 0.77 (.00) | 0.32 (.00) |
| Affluent Investor Dummy | | | | | 0.66 (.00) | 0.10 (.00) | | | | |
| # of stocks (Diversification measure) | | | | | | | 0.97 (.00) | 0.50 (.47) | | |
| Avg Correlation (Diversification measure) | | | | | | | | | 0.76 (.00) | 0.17 (.00) |

## Table 3.3

### Local Vs Non-Local Utility Stock Roundtrip trade performance

This table compares the gross return for Local Vs Non-Local utility roundtrip repurchase transactions (i.e., returns earned on the first sale transaction after repurchase). A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock. Once the gross return is computed for each roundtrip transaction, I convert it to a monthly measure. Panel A compares roundtrip transaction at the following levels: (1) Transaction Level, and (2) Household Level. The transaction level returns, are equally weighted returns for utility roundtrip in various categories. While, the household level involves an equally weighted aggregation at the household level, with each household return calculated as the value weighted gross return for repurchases in each category. Panel C classifies the households into the following categories: (1) Households that bought either local or non-local utility stocks, (2) Households that bought both, local and non-local utilities. The analysis done for the latter category is done using matched t-test at the household level. The *Nominal Returns* refer to the raw returns. The *Market Adjusted Abnormal Return* presented is the difference between the return on the roundtrip trade and the return on the market ($R_i$-$R_m$). The *Industry Adjusted Return* is the difference between the return on the roundtrip trade and the value weighted return earned by all stocks in the same industry, excluding the repurchased stock. The 49 industry classifications are obtained from Kenneth R. French's online data library. The *Jensen's alpha* is computed using the CAPM $[(R_i-R_f)-\beta(R_m-R_f)]$. The beta is estimated using the Scholes and Williams (1977) procedure over a time window of 250 trading days prior to each transaction. The CRSP value weighted NYSE/AMEX/NASDAQ is used to proxy for market returns and the 1 month t-bill rate is used as the risk free rate. The *Fama and French (1993) 3 factor abnormal return* is the difference between the observed and predicted return on a model that adds size and value factors to the CAPM $[(R_i-R_f)-\beta_1(R_m-R_f)-\beta_2(SMB)-\beta_3(HML)]$. The *Carhart (1997) 4 factor abnormal return* is the difference between the observed and predicted return on a model that adds a momentum factor to the Fama French 3 Factor Model $[(R_i-R_f)-\beta_1(R_m-R_f)-\beta_2(SMB)-\beta_3(HML)-\beta_4(WML)]$. Similarly the last model stated as *Carhart 4 factor plus Industry factor abnormal return* is the difference between the observed and predicted return on a model that adds an industry return factor to the Carhart Model $[(R_i-R_f)-\beta_1(R_m-R_f)-\beta_2(SMB)-\beta_3(HML)-\beta_4(WML)-\beta_5(IND)]$. The IND factor is the value weighted return earned by all utility stocks, excluding the utility stock being analyzed. The SMB, HML, WML and IND factors capture size, value, and momentum and industry effects respectively. The coefficient in each of the above factor models except for the CAPM were estimated using OLS regressions over 250 trading days prior to each repurchase. The p-values for all t-tests are presented in parenthesis.

| Panel A: Utility Roundtrip Returns at transaction and household level | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Roundtrip transaction Level | | | | | | Household Level | | | | | |
| Gross Return Measure | Local (1) | | Non-Local (2) | | Difference (1-2) | | Local (1) | | Non-Local (2) | | Difference (1-2) | |
| Nominal Returns | 0.49 | (.21) | 0.92 | (.29) | -0.43 | (.78) | 0.21 | (.17) | 1.17 | (.29) | -0.97 | (.54) |
| Market-Adj Returns | 0.32 | (.20) | 0.96 | (.30) | -0.64 | (.70) | 0.13 | (.17) | 1.23 | (.30) | -1.09 | (.52) |
| Industry-Adj Returns | 0.37 | (.22) | 0.96 | (.30) | -0.59 | (.72) | 0.15 | (.19) | 1.23 | (.30) | -1.08 | (.53) |
| Jensen's Alpha | 0.27 | (.18) | 0.97 | (.30) | -0.70 | (.67) | 0.12 | (.14) | 1.24 | (.30) | -1.13 | (.51) |
| 3 Factor Adj Returns | 0.44 | (.22) | 1.00 | (.30) | -0.56 | (.74) | 0.18 | (.20) | 1.28 | (.30) | -1.11 | (.53) |
| 4 Factor Adj Returns | 0.43 | (.22) | 1.01 | (.30) | -0.58 | (.74) | 0.18 | (.19) | 1.29 | (.30) | -1.11 | (.53) |
| 5 Factor Adj Returns | 0.45 | (.22) | 1.00 | (.30) | -0.56 | (.75) | 0.18 | (.20) | 1.28 | (.30) | -1.10 | (.54) |

| Panel B: Utility Roundtrip Returns across Household Classifications | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Only Local/Non-Local Houshold Level | | | | | | Partially Local Households Level | | | | | |
| Gross Return Measure | Local (1) | | Non-Local (2) | | Difference (1-2) | | Local (1) | | Non-Local (2) | | Difference (1-2) | |
| Nominal Returns | 0.31 | (.20) | 1.44 | (.30) | -1.13 | (.62) | 0.03 | (.08) | 0.01 | (.00) | 0.02 | (.37) |
| Market-Adj Returns | 0.20 | (.20) | 1.51 | (.30) | -1.31 | (.59) | 0.02 | (.31) | 0.00 | (.69) | 0.02 | (.37) |
| Industry-Adj Returns | 0.24 | (.21) | 1.52 | (.30) | -1.28 | (.59) | 0.02 | (.27) | 0.00 | (.80) | 0.02 | (.26) |
| Jensen's Alpha | 0.17 | (.17) | 1.53 | (.31) | -1.36 | (.58) | 0.02 | (.21) | 0.00 | (.25) | 0.02 | (.34) |
| 3 Factor Adj Returns | 0.28 | (.22) | 1.59 | (.30) | -1.31 | (.60) | 0.02 | (.24) | 0.00 | (.67) | 0.02 | (.30) |
| 4 Factor Adj Returns | 0.27 | (.21) | 1.59 | (.30) | -1.32 | (.60) | 0.02 | (.28) | 0.00 | (.77) | 0.02 | (.32) |
| 5 Factor Adj Returns | 0.28 | (.22) | 1.59 | (.30) | -1.31 | (.61) | 0.02 | (.36) | 0.00 | (.35) | 0.02 | (.28) |

## Table 3.4

## Non-S&P 500 Vs S&P 500 Local Utility Stock Roundtrip trade performance

This table compared the performance of Non S&P 500 and S&P 500 roundtrips, within the subset of local utility stocks. Except for a difference in comparison categories, the methodology is identical to that describes in Table 4. The p-values for all t-tests are presented in parenthesis.

| Panel A: Local Utility Roundtrip Returns at transaction and household level | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Roundtrip transaction Level | | | | | | Household Level | | | | | |
| Gross Return Measure | Non-S&P 500 (1) | | S&P 500 (2) | | Difference (1-2) | | Non-S&P 500 (1) | | S&P 500 (2) | | Difference (1-2) | |
| Nominal Returns | 1.18 | (.23) | 0.02 | (.00) | 1.16 | (.15) | 0.45 | (.19) | 0.01 | (.00) | 0.44 | (.13) |
| Market-Adj Returns | 0.78 | (.21) | 0.01 | (.19) | 0.77 | (.13) | 0.30 | (.17) | 0.00 | (.78) | 0.30 | (.10) |
| Industry-Adj Returns | 0.91 | (.22) | 0.01 | (.12) | 0.90 | (.14) | 0.34 | (.19) | 0.00 | (1.00) | 0.34 | (.12) |
| Jensen's Alpha | 0.65 | (.19) | 0.01 | (.03) | 0.64 | (.12) | 0.26 | (.15) | 0.00 | (.19) | 0.25 | (.09) |
| 3 Factor Adj Returns | 1.08 | (.23) | 0.01 | (.11) | 1.07 | (.15) | 0.41 | (.20) | 0.00 | (.81) | 0.41 | (.12) |
| 4 Factor Adj Returns | 1.05 | (.22) | 0.01 | (.13) | 1.04 | (.14) | 0.40 | (.19) | 0.00 | (.92) | 0.40 | (.12) |
| 5 Factor Adj Returns | 1.09 | (.23) | 0.00 | (.45) | 1.09 | (.14) | 0.42 | (.19) | 0.00 | (.22) | 0.42 | (.12) |
| Panel B: Local Utility Roundtrip Returns across Household Classifications | | | | | | | | | | | | |
| | Only Local Housholds Level | | | | | | Partially Local Households Level | | | | | |
| Gross Return Measure | Non-S&P 500 (1) | | S&P 500 (2) | | Difference (1-2) | | Non-S&P 500 (1) | | S&P 500 (2) | | Difference (1-2) | |
| Nominal Returns | 0.77 | (.21) | 0.01 | (.00) | 0.75 | (.12) | 0.05 | (.13) | 0.01 | (.05) | 0.04 | (.18) |
| Market-Adj Returns | 0.50 | (.20) | 0.00 | (.64) | 0.50 | (.11) | 0.04 | (.25) | 0.00 | (.31) | 0.05 | (.16) |
| Industry-Adj Returns | 0.59 | (.21) | 0.00 | (.53) | 0.59 | (.12) | 0.04 | (.21) | 0.00 | (.35) | 0.04 | (.13) |
| Jensen's Alpha | 0.43 | (.18) | 0.01 | (.08) | 0.42 | (.09) | 0.05 | (.21) | 0.00 | (.99) | 0.05 | (.17) |
| 3 Factor Adj Returns | 0.70 | (.22) | 0.00 | (.38) | 0.70 | (.12) | 0.04 | (.21) | 0.00 | (.53) | 0.04 | (.14) |
| 4 Factor Adj Returns | 0.68 | (.21) | 0.00 | (.40) | 0.68 | (.12) | 0.05 | (.24) | 0.00 | (.40) | 0.05 | (.16) |
| 5 Factor Adj Returns | 0.71 | (.22) | 0.00 | (.81) | 0.71 | (.12) | 0.05 | (.26) | -0.01 | (.05) | 0.06 | (.14) |

## Table 3.5

## Regional Vs Non-Regional Local Utility Stock Roundtrip trade performance

This table compared the performance of Regional (i.e. utility firms operating only in a single state) and Non-Regional (i.e. utility firms operating in multiple states) roundtrips, within the subset of local utility stocks. Except for a difference in comparison categories, the methodology is identical to that describes in Table 4. The p-values for all t-tests are presented in parenthesis.

| Panel A: Local Utility Roundtrip Returns at transaction and household level | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All Trades | | | | | | All Households | | | | | |
| Gross Return Measure | Regional (1) | | Non-Regional (2) | | Difference (1-2) | | Regional (1) | | Non-Regional (2) | | Difference (1-2) | |
| Nominal Returns | 0.94 | (.23) | 0.03 | (.00) | 0.91 | (.25) | 0.36 | (.19) | 0.02 | (.00) | 0.34 | (.22) |
| Market-Adj Returns | 0.62 | (.21) | 0.01 | (.08) | 0.60 | (.23) | 0.24 | (.18) | 0.01 | (.35) | 0.23 | (.20) |
| Industry-Adj Returns | 0.72 | (.22) | 0.02 | (.05) | 0.70 | (.24) | 0.27 | (.20) | 0.01 | (.31) | 0.27 | (.21) |
| Jensen's Alpha | 0.51 | (.20) | 0.02 | (.02) | 0.49 | (.22) | 0.20 | (.16) | 0.01 | (.11) | 0.19 | (.19) |
| 3 Factor Adj Returns | 0.85 | (.23) | 0.02 | (.03) | 0.83 | (.25) | 0.32 | (.21) | 0.01 | (.20) | 0.31 | (.23) |
| 4 Factor Adj Returns | 0.83 | (.23) | 0.02 | (.04) | 0.81 | (.24) | 0.32 | (.20) | 0.01 | (.26) | 0.31 | (.22) |
| 5 Factor Adj Returns | 0.86 | (.23) | 0.01 | (.11) | 0.85 | (.25) | 0.33 | (.20) | 0.00 | (.58) | 0.32 | (.22) |

| Panel B: Local Utility Roundtrip Returns across Household Classifications | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Only Local Housholds | | | | | | Partially Local Households | | | | | |
| Gross Return Measure | Regional (1) | | Non-Regional (2) | | Difference (1-2) | | Regional (1) | | Non-Regional (2) | | Difference (1-2) | |
| Nominal Returns | 0.57 | (.22) | 0.03 | (.02) | 0.55 | (.25) | 0.05 | (.14) | 0.01 | (.01) | 0.03 | (.26) |
| Market-Adj Returns | 0.37 | (.21) | 0.01 | (.30) | 0.36 | (.23) | 0.03 | (.28) | 0.00 | (.85) | 0.04 | (.27) |
| Industry-Adj Returns | 0.44 | (.22) | 0.01 | (.24) | 0.42 | (.24) | 0.03 | (.24) | 0.00 | (.68) | 0.03 | (.21) |
| Jensen's Alpha | 0.31 | (.19) | 0.02 | (.14) | 0.29 | (.23) | 0.04 | (.23) | 0.00 | (.51) | 0.04 | (.27) |
| 3 Factor Adj Returns | 0.51 | (.23) | 0.02 | (.21) | 0.50 | (.26) | 0.03 | (.25) | 0.00 | (.83) | 0.03 | (.26) |
| 4 Factor Adj Returns | 0.50 | (.22) | 0.01 | (.23) | 0.48 | (.25) | 0.04 | (.26) | 0.00 | (.94) | 0.04 | (.25) |
| 5 Factor Adj Returns | 0.52 | (.23) | 0.01 | (.38) | 0.51 | (.25) | 0.04 | (.30) | -0.01 | (.27) | 0.05 | (.24) |

## Table 3.6

### Correlation between returns on utility stocks and changes in utility prices

This table presents the correlations between monthly returns on utility stocks and monthly changes in utility prices across all the states For each state, I calculate the correlation between monthly percentage change in utility prices and monthly return on portfolio of companies operating in that state. Both, the equal and value weighting specifications are used for measuring each state's portfolio returns and the correlations with respect to each of these return measures are presented below. The significance test for mean correlations is based on single sided t-tests. The significance test for median correlations is based on the Wilcoxon signed rank test. The p-values are presented in parenthesis.

| Measure | Equally weighted Returns | | Value weighted Returns | |
|---|---|---|---|---|
| | Electricity Providers | | | |
| Mean Correlation | -0.0561 | (.00) | -0.0550 | (.00) |
| Median Correlation | -0.0754 | (.00) | -0.0634 | (.02) |
| | Gas Providers | | | |
| Mean Correlation | 0.0073 | (.64) | 0.0303 | (.06) |
| Median Correlation | 0.0074 | (.76) | 0.0348 | (.07) |

## Table 3.7

## Diversification and utility investors

This table shows estimates of various diversification measures across the following categories of utility investors: (1) Non-utility stock investors (i.e. investors who never invested in utility stocks), (2) Non-local utility investors (i.e. utility investors than never bought their local utility stock), and (3) Local utility investors (i.e. utility investors than bought their local utility stock at least once).The number of stocks (STK) measure is the number of stock in a portfolio. The WSQ (i.e. weight square) measure is the sum of the square of the proportion invested in each security. The IND measure is the number of industries across which the household diversifies.[40] The IWSQ (i.e. industry weight squared) measure is the sum of the square of the proportion invested in each industry. The average correlation measure presents the average correlation for all portfolios that have at least 2 stocks. For each of the above measures a monthly average for each household is calculated. Thereafter these measures are averaged across all households within various utility investor categories to obtain the estimates presented in this table. The F-Test tests for the equality of mean across the three investor categories. The p-values for all statistics are presented in parenthesis.

| Type of investor (household) | # of households | # of stocks (STK) | | Weight Sq WSQ | | # of Industries (IND) | | Ind Weight Sq (IWSQ) | | Average Correlation | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Non-utility stock investors (A) | 31075 | 2.85 | (.00) | 0.70 | (.00) | 1.99 | (.00) | 0.7728 | (.00) | 0.6054 | (.00) |
| Non-local utility investors (B) | 6051 | 5.32 | (.00) | 0.50 | (.00) | 3.14 | (.00) | 0.5939 | (.00) | 0.4140 | (.00) |
| Local utility investors (C) | 5135 | 6.03 | (.00) | 0.48 | (.00) | 3.24 | (.00) | 0.5930 | (.00) | 0.4146 | (.00) |
| Local Utility Investor - Non Local Unility Investor (C-B) | | 0.71 | (.00) | -0.02 | (.00) | 0.10 | (.00) | -0.0008 | (.85) | 0.0006 | (.92) |
| F - Test | | 1695 | (.00) | 2541 | (.00) | 2940 | (.00) | 2377 | (.00) | 1675 | (.00) |

[40] There are total number of 10 industry sectors that have been classified according to the definitions available Kenneth R. French’s online data library

**Table 4.1**

**Summary of OTC (over-the-counter) trades**

Panel A, B and C present a summary of OTC (over-the-counter) stock trades made at a discount brokerage house during the period Jan 1991-Nov 1996. Commission is calculated as commission paid by the value of the trade. Panel D presents a snapshot of the market statistics obtained from the OTCBB website (www.otcbb.com) on a randomly chosen date, March 8, 2009, at 7:40pm. The statistics in Panel D provide credibility to the data classified as OTC trades at the discount brokerage house.

| Variable | # | Minimum | 10th Pctl | 25th Pctl | Median | 75th Pctl | 90th Pctl | Maximum | Mean |
|---|---|---|---|---|---|---|---|---|---|
| Panel A: Average trade price of penny stocks | | | | | | | | | |
| Price ($) | 5307 | 0.000055 | 0.001 | 0.015 | 0.675 | 3.375 | 12.05 | 5332.86 | 8.468075 |
| Panel B: Purchase transactions | | | | | | | | | |
| Commission (%) | 42054 | 0.00% | 0.73% | 1.30% | 2.46% | 4.66% | 7.26% | 1303.33% | 3.85% |
| Trade size ($) | 42054 | 0.4 | 750 | 1527.5 | 3350 | 6875 | 14880 | 3150000 | 7469.02 |
| Price ($) | 42054 | 0.001 | 0.68 | 1.91 | 7.625 | 30.125 | 45 | 15750 | 44.47862 |
| Panel C: Sale transactions | | | | | | | | | |
| Commission (%) | 33888 | 0.00% | 0.56% | 1.13% | 2.41% | 7.80% | 100.00% | 125000.00% | 25.35% |
| Trade Size ($) | 33888 | 0.01 | 1.6 | 420 | 2850 | 7500 | 17500 | 3450000 | 8015.1 |
| Price ($) | 33888 | 0.00001 | 0.001 | 0.5312 | 7 | 29.375 | 44.25 | 9500 | 29.92349 |

| Panel D: Snapshot of trading activity on OTCBB as on March 8, 2009 | | | | | |
|---|---|---|---|---|---|
| Name | Symbol | Last Price ($) | Change ($) | % Change | Volume |
| Top (%) gainers | | | | | |
| PolyPacific International Inc. | PLYPF | 0.7500 | 0.7000 | 1400.00% | 250 |
| Genmed Holding Corp. | GENM | 1.0100 | 0.9300 | 1162.50% | 250 |
| Top (%) Losers | | | | | |
| Xenacare Holdings Inc. | XCHO | 0.0150 | -0.1150 | -88.46% | 4.8 k |
| Sunvesta Inc. | SVSA | 0.2000 | -0.5500 | -73.33% | 5.0 k |
| Top ($) Gainers | | | | | |
| Farmers & Merchants Bank | FMBL | 3300.0000 | 250.0000 | 8.20% | 70 |
| First National Bank Alaska | FBAK | 1610.0000 | 9.0000 | 0.56% | 47 |
| Top ($) Losers | | | | | |
| Mechanics Bank | MCHB | 12500.0000 | -500.0000 | -3.85% | 1 |
| Burke & Herbert Bank & Trust Company | BHRB | 1277.0000 | -23.0000 | -1.77% | 39 |
| Volume Actives | | | | | |
| Remote Dynamics Inc. | RMTD | 0.0001 | 0.0000 | 0.00% | 118.5 m |
| Phoenix Interests Inc. | PXIT | 0.0002 | 0.0000 | 0.00% | 79.16 m |

## Table 4.2

### Demographic Characteristics and proportion of Non-OTC/OTC investors

This table presents uni-variate results that show the relationship between demographic characteristics and proportion of Non-OTC/OTC investors. For all categorical variables (namely gender, investment experience, marital status, retirement status, age and net-worth categories), the percentages in the columns correspond to the percentage of Non-OTC and OTC investors in that particular category. The age categories in Panel E are defined as follows: (1) Young (less than 40), (2)Middle-aged (between 40 and 65), and (3)Old (over 65). The only continuous variable analyzed is the size to net-worth ratio (SNW) presented in Panel F. Size to Net worth ratio is the ratio of the average monthly investment in common stock divided by the self reported net worth of the client. The net-worth quartiles (wealth proxy) in Panel G are defined as follow (1) Q1: Net-worth less than $75,000 (2) Q2: Net-worth from $75,000 to $100,000 (3) Q3: Net-worth from $100,000 to $250,000 (4) Q4: Net-worth greater than $250,000. Further details for other demographic variables are provided in the text. The p-values for relevant statistic are in parenthesis.

Panel A: Gender

| Investor Category | Male | Female | # Obs - Row |
|---|---|---|---|
| Non-OTC Investor | 53.08% | 58.04% | 20355 |
| OTC Investor | 46.92% | 41.96% | 17569 |
| # Obs - Column | 33363 | 4561 | 37924 |
| Chi-sq Stat (p-value) | 39.67 (<.0001) | | |

Panel B: Investment Experience

| Investor Category | Low | High | # Obs - Row |
|---|---|---|---|
| Non-OTC Investor | 55.56% | 45.83% | 11333 |
| OTC Investor | 44.44% | 54.17% | 11499 |
| # Obs - Column | 8929 | 13903 | 22832 |
| Chi-sq Stat (p-value) | 108.37 (<.0001) | | |

Panel C: Marital Status

| Investor Category | Single | Married | # Obs - Row |
|---|---|---|---|
| Non-OTC Investor | 53.56% | 54.47% | 18057 |
| OTC Investor | 46.44% | 45.53% | 15506 |
| # Obs - Column | 24676 | 8887 | 33563 |
| Chi-sq Stat (p-value) | 2.20 (.1381) | | |

Panel D: Retirement Status

| Investor Category | Non-Retired | Retired | # Obs - Row |
|---|---|---|---|
| Non-OTC Investor | 50.15% | 53.42% | 11002 |
| OTC Investor | 49.85% | 46.58% | 9804 |
| # Obs - Column | 3432 | 17374 | 20806 |
| Chi-sq Stat (p-value) | 12.32 (.0004) | | |

Panel E: Age Categories

| Investor Category | Young | Middle | Old | # Obs - Row |
|---|---|---|---|---|
| Non-OTC Investor | 55.37% | 53.39% | 50.65% | 23439 |
| OTC Investor | 44.63% | 46.61% | 49.35% | 20176 |
| # Obs - Column | 15043 | 23361 | 5311 | 43615 |
| Chi-sq Stat (p-value) | 37.72 (<.0001) | | | |

Panel F: Size to Net-worth (SNW)

| Investor Category | SNW | |
|---|---|---|
| Non-OTC Investor (A) | 0.33 | (.00) |
| OTC Investor (B) | 0.57 | (.00) |
| Difference (A-B) | -0.24 | (.00) |

Panel G: Net Worth Categories

| Investor Category | Quartile 1 | Quartile 2 | Quartile 3 | Quartile 4 | # Obs - Row |
|---|---|---|---|---|---|
| Non-OTC Investor | 53.70% | 50.07% | 49.56% | 45.46% | 11660 |
| OTC Investor | 46.30% | 49.93% | 50.44% | 54.54% | 11870 |
| # Obs - Column | 5689 | 6795 | 4429 | 6617 | 23530 |
| Chi-sq Stat (p-value) | 84.23 (<.0001) | | | | |

## Table 4.3

## Logistic Regression Analysis for impact of demographic characteristics on decision to invest in OTC stocks

To obtain efficient estimates in the presence of missing values I implement the following two approaches that have explained in greater length in the methodology section: (1) Approach 1: Missing value dummy approach, and (2) Approach 2: Multiple Imputation approach. The following variables were dummy coded: Male, Experience, Age categories, Net Worth categories, Married dummy and Retired Dummy. A description of the variables is available in Table 4.2. The Global Chi-square test jointly tests all the regressions coefficients. The p-values for relevant statistic are in parenthesis.

| | Approach 1 | | | Approach 2 | | |
|---|---|---|---|---|---|---|
| Independent Variables | Point Estimates | | Odds Ratio | Point Estimates | | Odds Ratio |
| Intercept | -0.60 | (.00) | 0.55 | -0.49 | (.00) | 0.61 |
| Male | 0.18 | (.00) | 1.20 | 0.12 | (.00) | 1.12 |
| Experience (High Vs Low) | 0.34 | (.00) | 1.41 | 0.31 | (.00) | 1.36 |
| Age: | | | | | | |
| Young Vs Middle Age | -0.09 | (.00) | 0.91 | -0.09 | (.00) | 0.92 |
| Old Vs Middle Age | 0.08 | (.02) | 1.09 | 0.09 | (.00) | 1.09 |
| Net Worth: | | | | | | |
| Q2 Vs Q1 | 0.18 | (.00) | 1.20 | 0.05 | (.05) | 1.05 |
| Q3 Vs Q1 | 0.21 | (.00) | 1.24 | 0.05 | (.04) | 1.05 |
| Q4 Vs Q1 | 0.35 | (.00) | 1.42 | 0.23 | (.00) | 1.25 |
| Married Dummy | 0.03 | (.19) | 1.03 | 0.01 | (.64) | 1.01 |
| Retired Dummy | -0.10 | (.02) | 0.91 | -0.03 | (.33) | 0.97 |
| SNW (Size to Net Worth Ratio) | 0.25 | (.00) | 1.28 | 0.07 | (.00) | 1.07 |
| Missing Variable Dummy: | | | | | | |
| Gender | 0.18 | (.00) | | | | |
| Age | -0.01 | (.79) | | | | |
| Investing Experience | 0.13 | (.00) | | | | |
| Net Worth (or SNW) | 0.15 | (.00) | | | | |
| Marriage status | 0.09 | (.00) | | | | |
| Retirement Status | -0.06 | (.19) | | | | |
| Global Chi-square Test | 793.39 | (.00) | | 610.72 | (.00) | |

**Table 4.4**

**Robustness check: Logistic Regression Analysis for impact of demographic characteristics on decision to invest in OTC stocks**

The table below presents logistic regression results with a different classification for OTC and Non-OTC investors. Households that made a buy trade in at least one OTC stock are classified as OTC investors. The remaining households making buy trades in Non-OTC stocks are classified as Non-OTC investors. For further details refer to caption for table 4.3. The Global Chi-square test jointly tests all the regressions coefficients. The p-values for relevant statistic are in parenthesis.

| | Approach 1 | | | Approach 2 | | |
|---|---|---|---|---|---|---|
| Independent Variables | Point Estimates | | Odds Ratio | Point Estimates | | Odds Ratio |
| Intercept | -1.17 | (.00) | 0.31 | -1.27 | (.00) | 0.28 |
| Male | 0.04 | (.28) | 1.05 | 0.05 | (.40) | 1.05 |
| Experience (High Vs Low) | 0.30 | (.00) | 1.35 | 0.26 | (.00) | 1.30 |
| Age: | | | | | | |
| Young Vs Middle Age | -0.06 | (.06) | 0.94 | -0.05 | (.04) | 0.95 |
| Old Vs Middle Age | 0.06 | (.16) | 1.06 | 0.06 | (.26) | 1.06 |
| Net Worth: | | | | | | |
| Q2 Vs Q1 | 0.09 | (.04) | 1.09 | 0.03 | (.42) | 1.03 |
| Q3 Vs Q1 | 0.10 | (.03) | 1.11 | 0.07 | (.14) | 1.07 |
| Q4 Vs Q1 | 0.27 | (.00) | 1.31 | 0.27 | (.00) | 1.30 |
| Married Dummy | 0.02 | (.46) | 1.02 | 0.01 | (.74) | 1.01 |
| Retired Dummy | 0.01 | (.88) | 1.01 | 0.04 | (.53) | 1.04 |
| SNW (Size to Net Worth Ratio) | 0.11 | (.00) | 1.11 | 0.04 | (.00) | 1.04 |
| Missing Variable Dummy: | | | | | | |
| Gender | 0.12 | (.05) | | | | |
| Age | -0.06 | (.18) | | | | |
| Investing Experience | 0.05 | (.30) | | | | |
| Net Worth (or SNW) | 0.02 | (.71) | | | | |
| Marriage status | 0.07 | (.07) | | | | |
| Retirement Status | -0.01 | (.86) | | | | |
| Global Chi-square Test | 525.90 | (.00) | | 334.71 | (.00) | |

## Table 4.5

### Portfolio Characteristics and Non-OTC/OTC investors

This table presents the surveys portfolio characteristics for Non-OTC/OTC investors. The portfolio turnover in Panel A is calculated according to the methodology specified in Barber and Odean (2000). In Panel B, the Weight Square measure is the sum of the square of the proportion invested in each security. All the stocks are divided into 10 industries sectors as per the definitions available on Kenneth R. French's online data library. The Ind Weight Sq measure is the sum of the square of the proportion invested in each industry. Panel C presents logistic regression analysis for impact of demographic characteristics on decision to invest in otc stocks (yes or no). The definition of large cap and micro cap are provided in Table 4.6.

Panel A: Portfolio Turnover

| Category | Turnover | |
|---|---|---|
| Comparison of Non-otc stock portfolio turnover: | | |
| Only Non-otc stock Investors(A) | 6.24% | (.00) |
| Otc stock Investors (B) | 8.15% | (.00) |
| Difference (A-B) | -1.91% | (.00) |
| Paired t-test for difference in turnover among portfolios of OTC Investors: | | |
| Non-Otc Portfolio - Otc Portfolio | 3.11% | (.00) |

Panel B: Portfolio Diversification

| Investor Category | # of Stocks | | Weight Square | | # of Industries | | Ind Weight Sq | |
|---|---|---|---|---|---|---|---|---|
| Only Non-otc stock Investors (A) | 3.44 | (.00) | 0.64 | (.00) | 2.30 | (.00) | 0.72 | (.00) |
| Otc stock Investors (B) | 5.68 | (.00) | 0.51 | (.00) | 3.05 | (.00) | 0.62 | (.00) |
| Difference (A-B) | -2.23 | (.00) | 0.12 | (.00) | -0.76 | (.00) | 0.10 | (.00) |

Panel C: Logisitic regressions modelling otc stock investing decisions

| Independent Variables | Point Estimates | | Odds Ratio |
|---|---|---|---|
| Intercept | 0.20 | (.00) | |
| Non-otc stock turnover | 0.43 | (.00) | 1.53 |
| # of stocks | 0.02 | (.00) | 1.02 |
| Portfolio allocation (%) by stock cap: | | | |
| Large Cap | -0.09 | (.00) | 0.91 |
| Micro Cap | 0.02 | (.01) | 1.02 |
| Chi-square test | 3400 | (.00) | |

## Table 4.6

### Asset allocation and OTC stocks

This table depicts the distribution of investments across various Equity and Fixed Income instruments. Equities consist of following: (1) Stocks divided into large-cap (market cap > $ 10 billion), medium-cap ($ 2 billion<market cap ≤ $10 billion), small-cap ($300 million<market cap ≤ $2 billion), micro-cap (market cap ≤ $ 300 million) and over-the-counter stocks, (2) Equity Mutual Funds (includes Unit Investment Trusts). Fixed Income Instruments are divided into: (1) Bonds and (2) Fixed Income Mutual Funds (includes Unit Investment Trusts). The F-Test in Panel C tests for equality across all familiarity quintiles. All the averages presented in Panel A, B & C are significant at $\alpha$=.01

Panel A: Distribution across wealth categories

| Household | # of | Fixed Income | | Equity - Stocks | | | | | Equity | |
|---|---|---|---|---|---|---|---|---|---|---|
| Net Worth | households | Direct | Indirect | Large | Medium | Small | Micro | OTC | Indirect | Others |
| All Households | 23530 | 6.68% | 2.54% | 27.52% | 14.95% | 10.69% | 13.47% | 2.69% | 18.22% | 2.74% |
| 0-50 k | 4685 | 4.62% | 1.91% | 27.44% | 15.48% | 10.95% | 15.09% | 3.15% | 18.39% | 2.39% |
| 50 - 100 k | 7799 | 5.64% | 2.29% | 28.14% | 15.15% | 10.94% | 13.42% | 2.77% | 18.53% | 2.63% |
| 100 - 500 k | 8757 | 7.60% | 2.85% | 26.94% | 14.73% | 10.63% | 13.29% | 2.49% | 18.07% | 2.88% |
| 500 k - 1 m | 1616 | 10.89% | 3.72% | 27.25% | 14.16% | 9.67% | 11.15% | 2.16% | 17.45% | 3.29% |
| > 1m | 673 | 11.13% | 2.88% | 29.01% | 13.56% | 9.24% | 10.79% | 2.46% | 17.19% | 3.32% |

Panel B: Distribution across otc and non-otc investors

| Penny Group | # of households | Fixed Income Direct | Fixed Income Indirect | Equity - Stocks Large | Equity - Stocks Medium | Equity - Stocks Small | Equity - Stocks Micro | Equity - Stocks OTC | Equity Indirect |
|---|---|---|---|---|---|---|---|---|---|
| Non-OTC Investors (a) | 34724 | 5% | 2% | 34% | 18% | 12% | 14% | 0.00% | 12% |
| OTC Investors (b) | 30800 | 5% | 1% | 27% | 16% | 12% | 18% | 7.35% | 9% |
| Difference (b-a) | | 38.08 | 50.34 | 738 | 90.22 | 2.4 | 510.27 | 7668 | 312.79 |

Panel C: Distribution across otc quintiles

| OTC Invt Quintile | # of households | Fixed Income Direct | Fixed Income Indirect | Equity - Stocks Large | Equity - Stocks Medium | Equity - Stocks Small | Equity - Stocks Micro | Equity - Stocks OTC | Equity Indirect |
|---|---|---|---|---|---|---|---|---|---|
| 1 (Lowest) | 6160 | 4% | 1% | 34% | 20% | 13% | 17% | 0.06% | 7% |
| 2 | 6160 | 4% | 1% | 32% | 20% | 14% | 18% | 0.46% | 8% |
| 3 | 6160 | 5% | 1% | 28% | 18% | 13% | 21% | 1.71% | 8% |
| 4 | 6160 | 5% | 2% | 27% | 15% | 12% | 20% | 5.85% | 9% |
| 5 (Highest) | 6160 | 5% | 2% | 14% | 8% | 8% | 16% | 28.68% | 12% |
| F-test | | 4.75 | 13.06 | 546.19 | 480.59 | 163.73 | 64.35 | 7164.69 | 72.37 |

## Table 4.7

### Returns on round-trip trades across various OTC-quintiles

This table shows presents average roundtrip gross and net return earned on otc stock trades by households in each otc quintile. The net returns are obtained after adjusting gross returns for commission costs. A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock. For each household otc transaction value weighted net return is computed. Thereafter the value weighted returns are averaged across all households assigned to a familiarity quintile to arrive at the figure presented below. The *Small Adjusted Return* is the difference between the return on the roundtrip trade and the value weighted return earned by on the lowest market capitalization stocks in the CRSP database. The *Market Adjusted Abnormal Return* is the difference between the return on the roundtrip trade and the return on the market. The F-Test tests for equality across the otc quintiles. The p-values for all statistics are presented in parenthesis.

| Panel A: Gross Nominal and Abnormal Returns | | | | | | |
|---|---|---|---|---|---|---|
| OTC Invt | Gross Returns | | | | | |
| Quintile | Nominal | | Small Cap Adj | | Market Adjusted | |
| 1 (Lowest) | 11.77% | (.00) | -14.27% | (.00) | 1.80% | (.59) |
| 2 | 3.95% | (.00) | -14.82% | (.00) | -3.86% | (.00) |
| 3 | 9.27% | (.00) | -16.37% | (.00) | -0.63% | (.58) |
| 4 | 13.10% | (.00) | -19.75% | (.00) | 0.59% | (.69) |
| 5 (Highest) | 15.69% | (.00) | -25.44% | (.00) | 0.59% | (.78) |
| 5-1 (High-Low) | 3.91% | (.34) | -11.17% | (.00) | -1.20% | (.68) |
| F-test | 5.48 | (.00) | 1.08 | (.36) | 6.16 | (.00) |
| Panel B: Net Nominal and Abnormal Returns | | | | | | |
| OTC Invt | Net Returns | | | | | |
| Quintile | Nominal | | Small Cap Adj | | Market Adjusted | |
| 1 (Lowest) | 5.10% | (.10) | -20.95% | (.00) | -4.88% | (.11) |
| 2 | -1.06% | (.40) | -19.84% | (.00) | -8.88% | (.00) |
| 3 | 4.15% | (.00) | -21.49% | (.00) | -5.75% | (.00) |
| 4 | 7.82% | (.00) | -25.02% | (.00) | -4.69% | (.00) |
| 5 (Highest) | 9.89% | (.00) | -31.24% | (.00) | -5.20% | (.01) |
| 5-1 (High-Low) | 4.79% | (.20) | -10.29% | (.00) | -0.32% | (.86) |
| F-test | 5.65 | (.00) | 7.40 | (.00) | 0.99 | (.41) |

## Table 4.8

### Value-weighted returns on round-trip trades across price categories

This table presents the transaction value-weighted returns (gross and net) for all round-trip transactions in otc stocks across various price categories. The net returns are obtained after adjusting gross returns for commission costs. A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock. The *Small Adjusted Return* is the difference between the return on the roundtrip trade and the value weighted return earned by on the lowest market capitalization stocks in the CRSP database. The *Market Adjusted Abnormal Return* is the difference between the return on the roundtrip trade and the return on the market. *Avg Duration of Roundtrip* refers to the average number of days for roundtrips across price categories. The p-values for all statistics are presented in parenthesis.

Panel A: Gross Nominal & Abnormal Returns

| Price | # of | Avg Duration | Gross Returns | | | | | |
|---|---|---|---|---|---|---|---|---|
| Category | Roundtrips | of Roundtrip | Nominal | | Small Cap Adj | | Market Adj | |
| <= $1 | 1500 | 252 | 10.41% | (.17) | -6.40% | (.40) | 0.47% | (.95) |
| $1-$5 | 4701 | 255 | 7.81% | (.00) | -12.29% | (.00) | -0.59% | (.57) |
| $5-$10 | 2002 | 181 | -0.34% | (.72) | -13.94% | (.00) | -5.96% | (.00) |
| $10-20 | 3920 | 162 | 14.70% | (.00) | -0.48% | (.53) | 9.14% | (.00) |
| $20-$100 | 4219 | 197 | 1.98% | (.00) | -19.67% | (.00) | -5.05% | (.00) |
| > $100 | 2780 | 203 | -14.31% | (.00) | -31.63% | (.00) | -22.12% | (.00) |

Panel B: Net Nominal & Abnormal Returns

| Price | Commission | Net Returns | | | | | |
|---|---|---|---|---|---|---|---|
| Category | Costs | Nominal | | Small Cap Adj | | Market Adj | |
| <= $1 | 9.80% | 0.61% | (.93) | -16.20% | (.02) | -9.32% | (.18) |
| $1-$5 | 5.41% | 2.40% | (.01) | -17.70% | (.00) | -6.00% | (.00) |
| $5-$10 | 2.64% | -2.98% | (.00) | -16.59% | (.00) | -8.60% | (.00) |
| $10-20 | 2.19% | 12.51% | (.00) | -2.67% | (.00) | 6.95% | (.00) |
| $20-$100 | 1.49% | 0.49% | (.29) | -21.16% | (.00) | -6.54% | (.00) |
| > $100 | 0.75% | -15.06% | (.00) | -32.38% | (.00) | -22.86% | (.00) |

## Table 4.9

## Distribution of round-trip trade nominal returns

This table presents the distribution of gross and net nominal roundtrip trade returns in otc stocks across various price categories. A roundtrip transaction is defined as a buy and a subsequent sale of a particular stock.

| Panel A: Gross Return distribution | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Price | # of | Gross Return | | | | | | | | |
| Category | roundtrips | Skewness | Mean | Minimum | 10th percentile | 25th Percentile | Median | 75th Percentile | 90th Percentile | Maximum |
| <= $1 | 1500 | 19 | 39.45% | -99.96% | -82.21% | -58.21% | -17.27% | 38.74% | 107.75% | 11775.00% |
| $1-$5 | 4701 | 5 | 19.49% | -100.00% | -54.55% | -25.15% | 2.17% | 38.98% | 96.12% | 1892.13% |
| $5-$10 | 2002 | 3 | 5.61% | -99.98% | -50.45% | -20.00% | 1.01% | 22.58% | 56.86% | 516.50% |
| $10-20 | 3920 | 4 | 19.16% | -99.18% | -27.95% | -8.35% | 7.73% | 30.44% | 71.43% | 751.11% |
| $20-$100 | 4219 | 10 | 6.95% | -99.96% | -34.62% | -11.84% | 3.75% | 18.52% | 43.18% | 1590.32% |
| > $100 | 2780 | -1 | -4.94% | -89.84% | -71.52% | -8.66% | 4.07% | 15.99% | 31.70% | 141.50% |

| Panel B: Net Return distribution | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Price | # of | Net Return | | | | | | | | |
| Category | roundtrips | Skewness | Mean | Minimum | 10th percentile | 25th Percentile | Median | 75th Percentile | 90th Percentile | Maximum |
| <= $1 | 1500 | 19 | 22.59% | -100.00% | -86.33% | -64.36% | -27.96% | 23.05% | 86.00% | 10643.92% |
| $1-$5 | 4701 | 5 | 10.15% | -100.00% | -59.58% | -31.26% | -4.40% | 29.21% | 78.26% | 1604.62% |
| $5-$10 | 2002 | 2 | 0.42% | -100.00% | -54.06% | -24.00% | -2.63% | 17.30% | 49.35% | 462.51% |
| $10-20 | 3920 | 4 | 14.88% | -100.00% | -31.39% | -11.34% | 4.83% | 26.24% | 64.76% | 719.16% |
| $20-$100 | 4219 | 9 | 3.75% | -99.96% | -36.97% | -14.55% | 1.11% | 15.35% | 38.48% | 1422.86% |
| > $100 | 2780 | -1 | -7.38% | -95.52% | -72.17% | -11.45% | 2.17% | 12.69% | 27.97% | 140.28% |

**Figure 2.1**

**Net Market Adjusted returns on each roundtrip and Decision to Repurchase**

This above figure presents the net market adjusted returns earned in each roundtrip made by the average household in a particular stock. A roundtrip is defined as a buy and a subsequent sale of a particular stock. "Round trip number" is the serial roundtrip trade number assigned each time a stock is repurchased by a household. In this table each roundtrip is assigned to either of the two categories: (1) No further Roundtrips (2) More Roundtrips. "No further Roundtrips" is the last roundtrip after which a household does not conduct any further roundtrips in a particular stock. "More Roundtrips" is the roundtrip after which the household subsequently makes another roundtrip transaction in the same stock. The returns for each "Round Trip Number" are obtained by averaging the returns across all roundtrip trades with the same category.

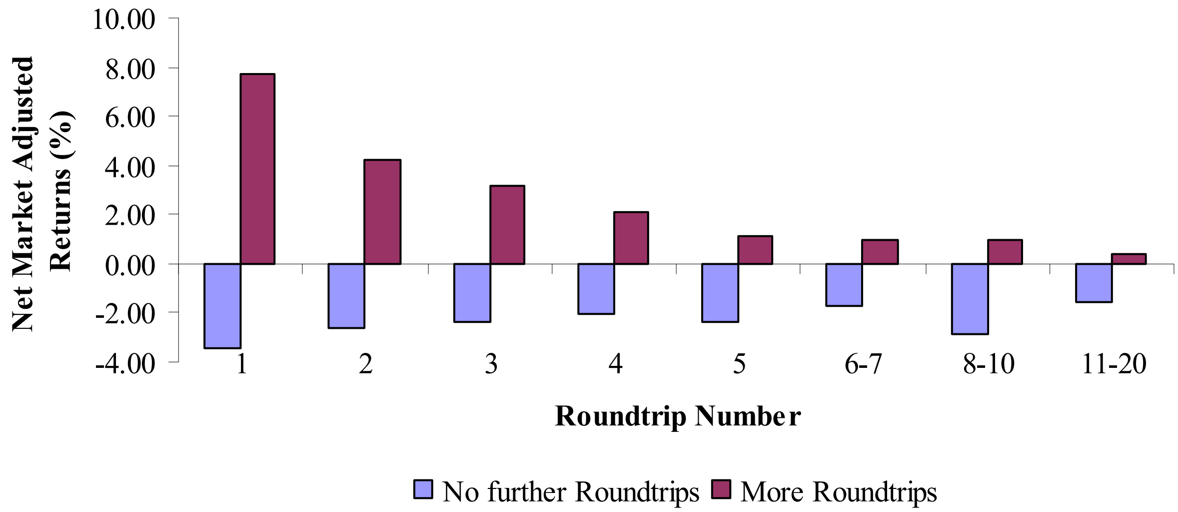

**Figure 2.2**

**Investment composition among various familiarity quintiles**

The above figure presents the breakup of the average monthly investment amongst common stocks, mutual funds and other instruments for the average households. For each household the average monthly investment in common stocks, mutual funds and other instruments is calculated. Thereafter for each familiarity quintile the average percentages for all households is calculated.

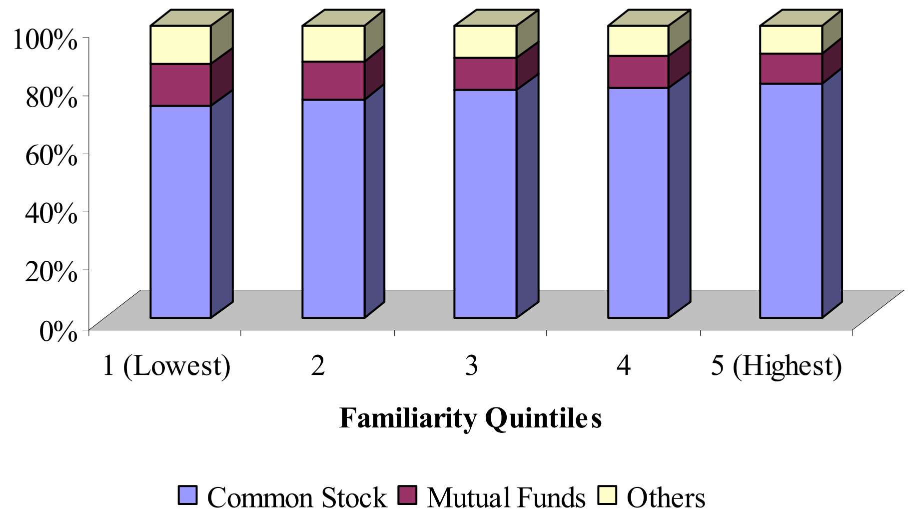

**Figure 3.1(a)**

**Distribution of number of purchases across various distance based categories**

The figures below present pie graphs of the distribution of number and value of purchase transactions in direct utility firms across various distance categories. Distance is determined as the distance between the household and the closest state centroid in which a firm operates.

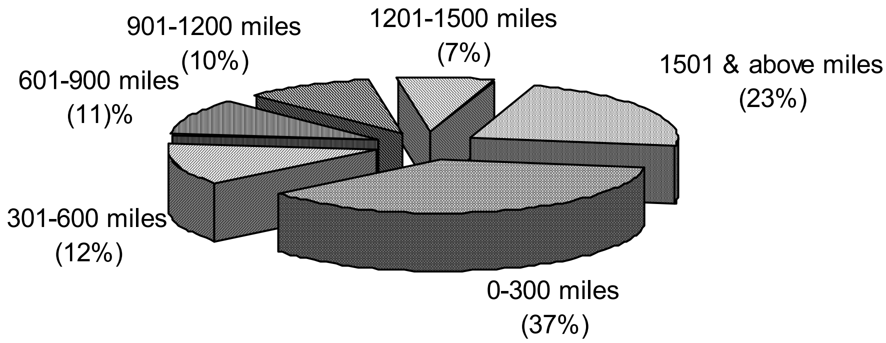


**Figure 3.1(b)**

**Distribution of value of purchases across various distance based categories**

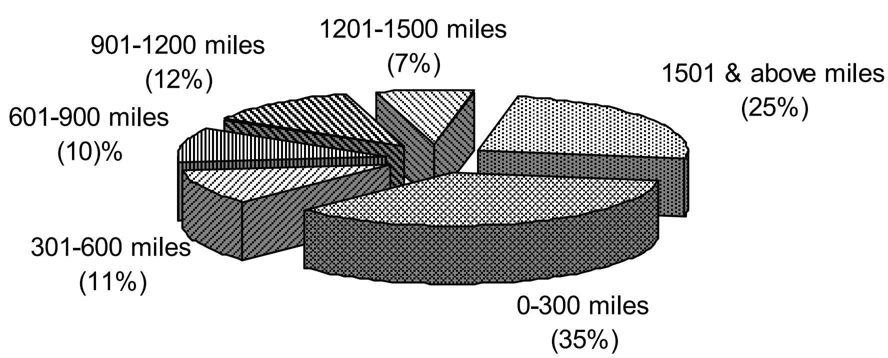

**Figure 4.1(a)**

**Average annualized returns in months prior to first purchase of an otc stock**

The portfolio return is the return earned on average return earned by households in their NYSE/NASDAQ/AMEX (non-otc) stock portfolio The market return is calculated using the CRSP value-weighted index. The Small Cap return refers to the returns earned on the lowest market capitalization decile stocks in the CRSP database.

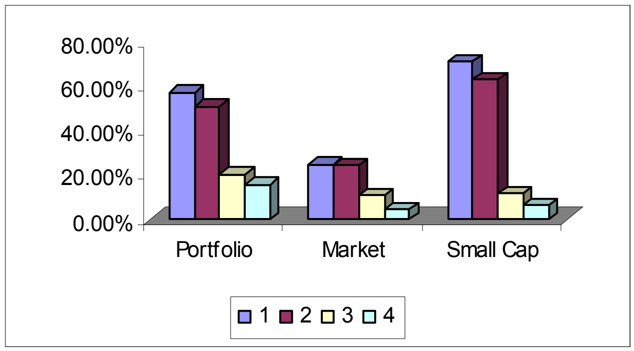


**Figure 4.1(b)**

**Average annualized returns in months prior to purchase of otc stocks**

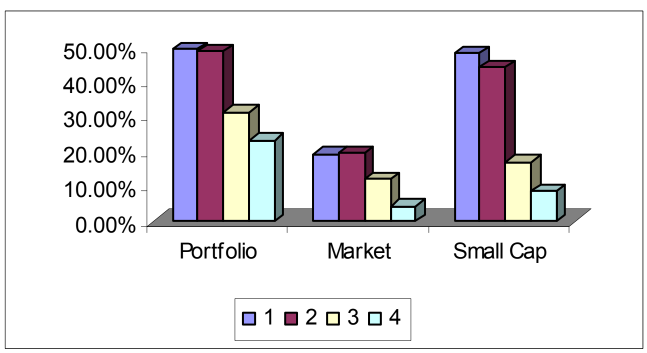

**Figure 4.1(c)**

**Average annualized returns in months prior to sale of otc stocks**