

Mark Twain's Cat: Industry Investment Experience, Categorical Thinking and Stock Selection*

Xing Huang[†]

Michigan State University - Department of Finance

October 2013

Abstract

This paper studies the impact of prior investment experience in an industry on the subsequent purchase of new stocks in the same industry. Using trading records data for households at a large discount broker from 1991 to 1996, I establish that the experience of positive excess returns in a given industry increases the probability of purchasing similar stocks in that industry relative to other industries. This result is robust to industry momentum, wealth effects, and investor heterogeneity. The effect decays when the experience is further in the past. Furthermore, I find that investor sophistication mitigates this experience effect. These results are consistent with mechanisms where investors put more weight on their own experience than on other available historical information when updating the beliefs about an industry's future return. The results are also consistent with investors learning about their stock-picking ability in an industry from their experienced outcomes.

*The cat, having sat upon a hot stove lid, will not sit upon a hot stove lid again. But he won't sit upon a cold stove lid, either. —Mark Twain

[†]Email: huangx@bus.msu.edu. I am grateful for very helpful suggestions and comments of Stefano DellaVigna, David Laibson, Ulrike Malmendier, Gregor Matvos, Terrance Odean, Andrei Shleifer, Adam Szeidl, Nancy Wallace and seminar participants at UC Berkeley (Finance, Financial Economics Lunch) and London Business School Trans-Atlantic Doctoral Conference. I would also like to thank an anonymous discount broker for providing data on individual investors' trades and Terrance Odean for his help in obtaining and understanding the dataset. All errors are my own.

1 Introduction

How do individuals select stocks? Do they incorporate all available historical information and update their beliefs in a Bayesian fashion? Or do they weigh differently their personal investment experiences and other statistical information? A growing body of evidence shows that past experience affects investors' choices in many financial decisions, including household risk taking, mutual fund investment style, and corporate financing decisions¹. Past experience may influence households' purchase decisions on common stocks as well. Barber et al. (2010) show that the effect of positive past investment experiences in a particular stock increases the likelihood of repurchasing the stock that was owned previously. But do past investment experiences also influence investors' propensity to buy other similar stocks? This question is important because it helps shed light on how investors update their beliefs about stock returns.

In this paper, I explore whether investors put more weight on their idiosyncratic personal experience of investment in an industry when they make decisions about purchasing new stocks. Consider, for example, two individual investors, A and B, who form portfolios in 1991. Investors A and B both invest in the insurance industry, but they invest in different companies. A picks Pioneer Financial Services Inc. while B picks Conseco Inc. One year later, investor A has a paper loss of 30% whereas investor B's investment doubles. Given the idiosyncratic realization of returns, after enjoying a huge gain from investing in insurance industry, will B, when compared to A, be more prone to invest in *other* stocks in that industry? The psychology literature suggests that personally experienced outcomes have a greater impact on personal decisions than information acquired merely by reading, which comes without personal involvement (Weber, Bockenholt, Hilton, and Wallace, 1993; Hertwig, Barron, Weber, and Erev, 2004). The experience hypothesis predicts that in an industry where investors have prior investment experience, it is more likely for them to buy new stocks after they have experienced good rather than bad returns.

This paper exploits data from detailed trading records for households at a large discount

¹Barber, Odean, and Strahilevitz (2010), Chan, Chen, and Lakonishok (2002), Choi, Laibson, Madrian, and Metrick (2009), Graham and Narasimhan (2004), Kaustia and Knupfer (2008), Malmendier and Nagel (2011), Malmendier and Tate (2005), Malmendier, Tate, and Yan (2011)

broker from 1991 to 1996 (Barber and Odean, 2000) as a measure of investors' personal investment experiences, and explores whether the past experiences of these investors affect their subsequent purchase decisions at the industry level. The results indicate that the likelihood of investors purchasing new stocks in an experienced industry increases with their experienced excess return. Furthermore, the effect of experienced outcomes becomes weaker for the purchase decisions on stocks in more different industries. Specifically, experience in one industry influences the purchases in a similar industry by a smaller amount, and has almost no effect on the purchases in a dissimilar industry.

While these results are consistent with the view that past experience influences stock selection, there are four alternative mechanisms that may also drive the correlations between the past experiences and the future purchases in the same industry. First, momentum traders (Hong and Stein, 1999; DeLong, Shleifer, Summers, and Waldmann, 1990; Barberis and Shleifer, 2003) may be more likely to purchase new stocks in an industry that performed well in the past, regardless of their personal investment experience in that industry. To control for this momentum effect, I include past industry average returns in the regressions and show that my results are robust. A second alternative story is a wealth effect: a high experienced return may increase wealth, generating new purchases in all industries. The results mentioned above - that the experience effect spills overs only slightly to other-industry purchases - do not support this story. I also show that my results are robust to controls measuring the change in the investor's portfolio value. Third, an investor with a high investment ability is more likely to gain high returns and in the meantime may also be more active, which could drive the positive relation between experienced outcomes and future purchases. I measure investment ability by three trading characteristics variables and also test how the correlation varies with different degrees of sophistication. The evidence does not support this story either. A fourth relevant story is portfolio rebalancing. But because households should decrease their holdings in the industry that earns relatively higher returns in the past to rebalance their portfolio, this story actually leads in the direction opposite to the prediction of the experience hypothesis.

After controlling for industry momentum and time-varying investor heterogeneity, the results show that investors with positive market-adjusted returns in one industry have 1.86

pp higher propensity to buy new stocks in the same industry as opposed to those who earn negative market-adjusted returns. This magnitude corresponds to 16.60 percent of the average probability of 11.18 pp for purchasing new stocks in an industry. However, this effect that results from the experience of positive excess returns drops significantly with regard to the purchase of new stocks outside of the industry in which the investor has experience. The magnitude of this effect on the purchase of new stocks in an industry that is the most similar to the experienced industry is only one third of that for the purchase of new stocks in the original industry, whereas the effect on purchasing new stocks in an industry that differ significantly from the experienced industry is negligible.

I also test for the long-term effect of lagged experience outcomes. Because the sample only extends from 1991 to 1996, the longest horizon is chosen to be half of that sample period, i.e., three years. The results indicate that more recent experience has a stronger influence on purchase of new stocks. The effect of experience outcomes drops dramatically if the experience happens more than one year before the month of purchase. [Gallagher \(2012\)](#) finds a similar effect in the context of flood insurance; after a flood, the take-up rate steadily decreases in the flooded communities.

If the experience hypothesis is driving the relation between experienced outcomes and future purchases, the magnitude and significance of the results should vary with investor sophistication and portfolio diversification. As investors become more sophisticated, their past experiences may cancel each other, with the result that more recent experiences have less influence. I divide the households in the sample into four subgroups according to their self-reported investment sophistication as provided in the dataset and test the effects of experienced outcomes separately for each subgroup. The results indicate that the effect is most pronounced for the group with no experience or knowledge in investing, but insignificant for the group with extensive investment experiences, which provides further evidence in support of the experience hypothesis.

Furthermore, investors with more diversified portfolios may care less about the individual performance of each component; consequently, the experienced outcome on the industry level would have less effect. I also explore the variation in portfolio diversification by adding an interaction term between the experience variables and measures for portfolio diversification

in the regression. As predicted above, the results indicate that influence of past experiences declines with the diversification of the household portfolio.

The evidence that past investment experiences have an impact on future purchases at the industry level suggests that investors may apply categorical thinking in their stock investments. To gain additional insight into the operation of categorical thinking in conjunction with the experience effect, I investigate how the latter experience effect varies with the refinement of categories. I create a proxy to measure the ability of the experienced stocks to represent other stocks in the same industry. The results indicate that the experience effect is stronger when the experienced stocks are more representative. I also test within the same industry to determine whether the effect of experience in a given subindustry would spill over to another subindustry. The results do not suggest that investors have finer categorical thinking than Fama-French 10-industry classification.

This paper is related to several strands of literature. First, it contributes to the growing literature on investor behavior; specifically, I address how investors choose which stocks to purchase. Some papers focus on investors' cross sectional preferences of stocks,² while others discuss investors' stock purchase in a time series that connects buying decisions with past investment experience.³ This paper exploits the latter approach and emphasizes the affects of past experience on future purchase of new stocks at different levels of categorization. [Campbell, Ramadorai, and Ranish \(2013\)](#) also study the influences of investors past experiences on other behaviors, such as trading frequency, disposition bias and diversification. They find that turnover, disposition bias and underdiversification diminish when investors experience poor returns resulting from them.

The literature also indicates that experience affects a number of other financial decisions, including IPO subscriptions ([Kaustia and Knupfer, 2008](#)), 401(k) accounts portfolios ([Choi](#)

²For example, [Barber and Odean \(2007\)](#) study the attention-grabbing stocks and find that individual investors are net buyers of attention-grabbing stocks. [French and Poterba \(1991\)](#) and a lot of other papers document the home bias puzzle that individuals and institutions in most countries hold modest amounts of foreign equity even though observed returns on national equity portfolios suggest substantial benefits from international diversification.

³For example, recent researches find that investors tend to buy stocks with strong recent performance ([Odean, 1999](#); [Barber, Odean, and Zhu, 2009](#); [Jackson, 2003](#)). [Barber et al. \(2010\)](#) establish the trading patterns in more details about how investors' repurchase of stocks previously sold is affected by their investment experience in those stocks.

et al., 2009), stock market participation (Malmendier and Nagel, 2011), the investment style of fund managers (Chan et al., 2002), corporate external financing (Malmendier and Tate, 2005; Malmendier et al., 2011), etc. Even though the tests employed in this study are based on the idiosyncratic personal experiences of investors, an examination of the variation in experiences across cohorts (as in Malmendier and Nagel (2011)) may lead to similar results. Specifically, assume investors begin investing in the stock market in their mid twenties. For households whose members are between the ages of 30 and 40, I compute experiences as the average of industry returns during the preceding 10 years. Likewise, for the households whose members between the ages of 50 and 60, experiences are measured by the average of industry returns during the preceding 30 years. Figure 1 plots the difference in the fraction of households investing in an industry between above age cohorts against the difference of the experience in the same two groups. The figure suggests that more investors hold the stocks of an industry when their experience in that industry is better, which is consistent with the experience hypothesis.

This paper also contributes to the literature’s examination of investors’ categorical thinking or category learning behavior. Barberis and Shleifer (2003) assume that investors categorize risky assets into different styles and trade among styles depending on their relative performance, which derives excessive comovement of assets in the same style, but little comovement of assets in different styles and other asset pricing predictions. Peng and Xiong (2006) provide justification of category-learning behavior when attention is a scarce cognitive resource; they also generates features in return comovement and other predictions. The present paper provides empirical evidence that investors extrapolate their experience of stocks in one industry to their decision about similar stocks in that industry, and these results provide microfoundation for the effect of experience under categorical thinking or learning.

Section 2 describes the datasets used in the paper as well as the methodology, and presents summary statistics. Section 3 details the results for the examination of the influence of experience on the future purchase of new stocks. Section 4 explores the underlying mechanisms through an examination of how this experience effect varies with investor sophistication, portfolio diversification and categorization. Section 5 concludes.

2 Datasets and Methodology

2.1 Data Description

The dataset used in this paper includes the trading records of 78,000 households at a large discount brokerage house over the period 1991-1996. This dataset is used by [Barber and Odean \(2000\)](#) and others. Each household has at least one account, but some have many. I combine the trades of accounts within the same household and build observations at the household level. This paper only focus on investors' direct investments on common stock, so I exclude their investments in mutual funds, American depository receipts(ADRs), warrants, and options. The sample is further refined by removing observations with errors in trading records, short selling trades, etc. The number of households in the final sample is 47,993. More details about the restrictions I impose to select the sample for analysis are listed in the appendix. I use the Center for Research in Security Prices (CRSP) database to obtain information on stock prices for calculating investor experienced return or portfolio related variables.

One feature of this paper is to discuss investors' stock selection choices among a manageable number of categories, more specifically, industries. The Standard Industrial Classifications (SIC) codes are obtained from two resources: CRSP and Compustat.⁴ For most of the tests, stocks are classified into 10 industry groupings based on their SIC code according to a algorithm devised by [Fama and French \(1997\)](#). The 10 industry groupings include (1) consumer nondurables, (2) consumer durables, (3) manufacturing, (4) oil, gas, and coal extraction and products, (5) high technology, (6) telephone and television transmission, (7) wholesale, retail, and some services, (8) healthcare, medical equipment, and drugs, (9) utilities, and (10) others. I also exploit the Fama-French 48-industry classification in robustness tests and to define subindustries in further analysis.

⁴The first step is to match the Cusip of the stocks invested by the households with corresponding SIC code in CRSP. If a corresponding SIC code can not be matched in CRSP, a second round match is proceeded in Compustat.

2.2 Investors' Investment Experiences and Purchase Decisions

I construct investors' experienced returns in each industry by their trading records. I select a fixed window to measure experienced returns. For example, the experience window spans from the beginning to the end of each year. Note that the returns in a fixed window could be either realized or not. I do not use realized returns to measure experiences, because I want to avoid introducing any potential endogeneity.

The measures of experienced outcomes build on the market-adjusted experienced returns. For every household h , I denote by r_{hjt} the annualized return (either realized or not) of the stock j during the experience window t . The experienced returns of industry I , r_{hIt} , is defined as value-weighted average returns of stocks belonging to the industry ($j \in I$):

$$r_{hIt} = \frac{\sum_{j \in I} x_{hjt} r_{hjt}}{\sum_{j \in I} x_{hjt}} \quad (1)$$

where x_{hjt} is the dollar value allocated on stock j at the date of purchase, or at the beginning of the window if the date of purchase is before the window starts. The market-adjusted experienced return in industry I , er_{hIt}^m , is the difference between r_{hIt} and the market return of period t , R_{mt} :⁵

$$er_{hIt}^m = r_{hIt} - R_{mt} \quad (2)$$

I consider three measures of investors' past experienced outcome, which are three indicator variables denoting good, top and bottom experiences. The dummy for good experience, $Goodexp_{hIt}$, equals to one if household h earns a positive market-adjusted return in industry I during period t , i.e. $er_{hIt}^m > 0$. The dummy for top (bottom) experience, $Topexp_{hIt}$ ($Bottomexp_{hIt}$), equals one if er_{hIt}^m is above the 90th (below the 10th) percentile of market-adjusted experienced returns by all households during period t . While these indicator variables are all based on market-adjusted experienced returns which are relative measures of experience, the results are robust to other measures, such as those based on the raw level of experienced returns.

Investors' purchase decisions are measured by an indicator variable $B_{hI,t+1}^{new}$, which takes

⁵I will use lowercase letters to denote the experienced returns of households, such as r_{hIt} , er_{hIt}^m ; and uppercase letters to denote market and industry average returns, such as R_{mt} , R_{It} .

a value of one if household h purchases new stocks (those not previously owned in the experience window) in industry I in decision period $t + 1$ following the experience period t . Note that the purchase decisions only focus on new stocks, but not on previously owned stocks. If I do not exclude the previously owned stocks, the effect of past experiences on the industry level may be confounded with the effect on the experienced stocks themselves. [Barber et al. \(2010\)](#) find that investors are more likely to repurchase the stocks which have been previously sold for a gain; they also find investors prefer to purchase additional shares of stocks that have lost value since being purchased. Since the experienced return (previously defined) could be either realized or not, the effect on previously owned stocks can push the results either way. Therefore, to separately identify the effect on the industry level, I will only consider purchases of new stocks in the following tests.

2.3 Summary Statistics

Table 1 represents the summary statistics. Panel A reports the frequencies of the trades of buying new stocks and repurchasing previously owned stocks through the years of the decision periods (1992-1996). The purchases of new stocks account for a large portion (about 85%) of investors' overall purchase decisions.

Panel B and Panel C summarize statistics related to past experiences. Panel B reports the distribution of households' experience across industries for each year of experience window. First, there are relatively more households trading in some of the industries, such as (3) manufacturing, (5) high technology, (7) wholesale, and (8) health care, which attract households over 10%. But overall, households' participation in each industry is roughly balanced, which could help rule out the possibility that the results are driven by a concentration of trades in some particular industry. Second, this table could also show the distribution of households' experience among the industries is stable across years.

Panel C provides a first look at the distribution of experience outcomes within each industry for each year of experience window. We can observe both the cross sectional and the time series variations. For example, during 1991, a great portion of the households (81.2%) had bad experiences in the energy industry, while over half of the households had good experiences in other industries such as wholesale, health care and utilities. However,

households in the energy industry do not always have bad experiences. In 1993, over half of the households investing in energy had good experiences.

3 Industry Investment Experience and Stock Selection

In this section, I will study the effects of experienced returns in one industry on the decision to purchase new stocks in the same industry. I estimate a baseline specification using a probit model, and then present graphical evidence, followed by several robustness tests, such as using different measures of experienced outcomes, different industry classifications, addressing alternative explanations, etc. Finally, I examine the long term effect of prior experience outcomes on stock selection.

3.1 Graphical Evidence

The monotone increasing relationship between prior experienced outcomes and future purchases can be also illustrated as in Figure 2. I divide all the industry experienced outcomes into five bins, ordered by market-adjusted returns in each industry of each period. The right-most bars (group 5) correspond to the top 20% experienced returns, while the left-most bars (group 1) correspond to the bottom 20% experienced returns. The black bars in both figures represent the probability of buying new stocks in the experienced industry. Figure 2(a) employs the original data. It shows a roughly monotone increasing relationship: as experienced returns go up from quintile 1 through 5, investors are more likely to buy new stocks in the experienced industry, especially so in the upper tail of the experienced return. Figure 2(b) plots the average of generalized residuals within each quintile. The residuals are obtained from a probit model of regressing purchases of new stocks in one industry on control variables. The details about control variables will be discussed later. After removing the effects of controls (industry average, wealth effect and etc), the monotone increasing relationship becomes more striking.

3.2 Baseline Model

I start by modelling the probability of purchasing new stocks in one industry with a probit model. The dependent variable ($B_{hI,t+1}^{new}$) indicates the purchases of new stocks (those

not previously owned in period t) in industry I of period $t + 1$. The specification is written as follows:

$$P(B_{hI,t+1}^{New} = 1) = \Phi(\beta_0 + \beta_1 Exp_{hIt} + \beta_2 Goodexp_{hIt} + \beta_3 Topexp_{hIt} + \beta_4 Bottomexp_{hIt} + \Gamma' X_{hIt}) \quad (3)$$

where $\Phi(\cdot)$ denotes the cumulative standard normal distribution function. Exp_{hIt} is a dummy variable equal to one if household h has investment in industry I in period t . In the model, each observation corresponds to a household/industry/year pair. Because I use the Fama-French 10-industry classification, each household/year pair corresponds to 10 observations. The primary coefficients of interest are on the variables of experienced outcomes, which are measured by three dummy variables, $Goodexp_{hIt}$, $Topexp_{hIt}$, $Bottomexp_{hIt}$, equal to one if household h earns positive, above the 90th percentile, below the 10th percentile market-adjusted returns in industry I of period t . The experience hypothesis that investors overweight their experienced outcomes predicts a monotone positive relationship between experienced outcomes and future purchases, i.e. $\beta_2 > 0$, $\beta_3 > 0$, $\beta_4 < 0$.

3.2.A Control Variables

These predications may be consistent with other explanations as well, such as (1) industry momentum trading, (2) wealth effects, and (3) investors' heterogeneity. I include a vector of controls (X_{hIt}) to address these issues:

Industry Momentum Trading. Moskowitz and Grinblatt (1999) find a strong and prevalent momentum effect in the industry component of stock returns, which provides a way for investors to conduct momentum trading on industry level. Industry-level momentum trading could also lead to a positive relationship between prior experienced outcome and subsequent purchases. To control for this confounding effect, I include three industry average variables. They are all based on market-adjusted industry average return, $ER_{It} = R_{It} - R_{mt}$, where R_{It} denotes the value-weighted average return of industry I during period t . Corresponding to the dummy variables measuring experienced outcomes, the three industry average variables are created as: (1) $Goodind_{It}$: an indicator of industries with positive market-adjusted industry average returns during the experience window, i.e. $ER_{It} > 0$;

(2) $Topind_{It}$: an indicator of the industry with the highest market-adjusted industry average return; (3) $Bottomind_{It}$: an indicator of the industry with the lowest market-adjusted industry average return.

Wealth Effect. Investors with good experiences in some industries are more likely to have increases in their overall stock portfolios. If investors tend to purchase new stocks in all the industries when their overall portfolios earn profits, the positive correlation between good experiences and future purchases in a specific industry may hence show up. To address this explanation, I include a dummy variable indicating the overall value of the household's portfolio on common stocks increases during the experience window.

Investors' Heterogeneity. Investors differ in their investment ability and level of expertise. Some investors with superior ability may be better at picking misvalued securities or predicting economic prospects. These investors are more likely to gain high returns and have good experiences. Even though the new purchases are made randomly across industries, we may observe the positive correlation between good past experiences and future purchases. To ensure that investment ability is not driving the relation, I create three variables measuring trading characteristics as a proxy of investor's investment ability. An investor with higher investment ability may trade more frequently, own a larger portfolio or hold a greater number of stocks in the portfolio. The three variables calculated by using the beginning-of-month position data from the experience window are: (1) Average number of stocks in the beginning-of-month portfolios; (2) A logarithm of the average size of beginning-of-month portfolios; (3) A logarithm of the average of monthly turnover rate calculated following [Barber and Odean \(2000\)](#).⁶

3.2.B Basic Results

Table 2 presents the results. In addition the controls mentioned above, all specifications include year effects and industry effects. The standard errors are clustered by industry-year

⁶In each month during the sample period, I identify the common stocks held by each household at the beginning of month t from their position statements. To calculate monthly sales turnover, I match these positions to sales during month t . The monthly sales turnover is calculated as the shares sold times the beginning-of-month price per share divided by the total beginning-of-month market value of the household's portfolio. To calculate monthly purchase turnover, I match these positions to purchases during month $t - 1$. The monthly purchase turnover is calculated as the shares purchased times the beginning-of-month price per share divided by the total beginning-of-month market value of the portfolio. Finally, monthly turnover is calculated by averaging monthly sales turnover and monthly purchase turnover.

level. I report the marginal effect of each variable in the table.

The results for experience-related variables are consistent across all specifications. I find a significant and monotone increasing relationship between households' experienced outcomes and their future purchases within the same industry. Columns (1)-(4) include all three experienced-related variables and divide investors' experienced outcomes into four categories: (1) bottom; (2) bad but not bottom (base category); (3) good but not top; (4) top. According to Column (1), the propensity of buying new stocks in an industry is significantly 1.15 percentage points (pp) higher when the investor enjoys a good but not top experience relative to bad but not bottom. The propensity increases by 1.93 pp more when the investor earns a top experience. In addition, having a bottom experience makes the investor even more reluctant to purchase again in the same industry, the propensity drops by 1.29 pp compared to the base category. These results are consistent with the experience hypothesis that investors have higher propensity to purchase new stocks in the same industry if they experience higher returns in their past investment.

To better understand the economic magnitude, I only include the dummy indicating good experience in Column (5). Relative to having a bad experience, the investor with a good experience has a 1.86 pp higher probability to buy new stocks in that industry. This magnitude corresponds to 16.60 percent increase ($= 1.86 \text{ pp}/11.18 \text{ pp}$) of the probability if we normalize it by the average probability of buying new stocks in an industry within a single period (11.18 pp).⁷

As a comparison with the effect of personal experiences, the effect of industry average returns is quite different and exhibits a U-shaped relationship. The non-extreme industry average variable does not have a significant influence. but the industries with the highest and lowest market-adjusted return both have positive impacts on households' future purchases.

⁷In addition, I want to point out that the setting of baseline model is actually estimating a lower bound of the effect. If the investors are sorted into some industry they think they have information advantage or they are more familiar with, their experienced outcomes may have less influence. And as more and more experience the investors get in some industry, the effect of new experience tends to decrease. Ideally, I would like identify the industries that investors are not sorted into (for example, the investors are attracted by some exogenous events and then start to invest in some industry) and be able to observe investors' trading records since the first time they enter that industry. Therefore, given the setting of current dataset, the estimated magnitude of experience effect may be dampened due to sorting or diminishing influence as investors cumulate more and more experience in a certain industry.

This evidence is consistent with the “attention grabbing” effect found in [Barber and Odean \(2007\)](#). Individual investors tend to purchase stocks in the industries which exhibit big price moves, because the stocks in those industries catch their attention. The results of the wealth effect and trading characteristics controls are reported in Column (4). Consistent with our expectations, investors have a higher probability of buying new stocks when they earned money in the past year, trade more frequently and have a larger portfolio invested in the stock market.

As a side note, the marginal effect of the experience dummy is also significantly positive. It implies a significantly positive unconditional effect of prior investment experience. In other words, regardless of outcomes, personal involvement, on average, has a positive effect on the probability of future purchases in that industry. Several explanations could explain this effect. For example, the involvement catches investors’ attention, therefore, the stocks in this industry are more likely to enter the choice set for future purchase; or investors are sorted into certain industries because they may have worked in those industries, have information advantages and tend to buy stocks in those industries. The experience dummy is not the focus of this paper, but it is important to put it as a control in the regression.

3.3 Robustness Tests

These results are robust to alternative experience measures, subsample, explanations and industry classification. [Table 3](#) presents a series of robustness tests.

Instead of using dummy variables to measure experienced outcomes and industry average returns, Panel A directly applies the level of market-adjusted experienced returns and industry average returns. The coefficient on the experienced return is significant and positive, confirming that investors have higher propensity to buy new stocks in an industry if they earned higher returns in the same industry in the past. Different from the effect of past experience, the effect of industry average return is nonlinear and represents a U shape, which is also consistent with the evidence found in [Table 2](#).

In Panel B, I drop the industry-year observations if the household does not have investment in the industry during the past period. Then I run specification [\(3\)](#) without the experience dummy. As Panel B of [Table 3](#) shows, the results remain the same, indicating a

monotone increasing relationship between experienced outcomes and future purchases within the same industry.

Panel C considers an alternative story of mental accounting. If investors apply mental accounting (Thaler (1999)) and regard each industry as an account, they may rebalance portfolios only within an industry but not across industries. Such investors purchase stocks in an industry with the money from recent sales in that industry. Because sales are more likely to happen following good experiences due to disposition effect (Odean (1998)), purchases following the sales are more likely to happen as well. To address this issue, I exclude the purchases that happen within 30 days after the most recent sales in the same industry. In other words, the dependent variable is a dummy variable which indicates purchases of new stocks without recent sales. The results are virtually unchanged, though the magnitude does drop slightly. According to Column (5), good experience increases the propensity of buying new stocks without recent sales by 0.92 pp relative to bad experience, which corresponds to an increase of 11.35 percent ($= 0.92 \text{ pp}/8.08 \text{ pp}$) if normalized by the average probability of buying new stocks without recent sales in an industry.

I also exploit the Fama-French 48 industry classification as a robustness check. To make the definition of the top and bottom industries match with the definition of the top and bottom experiences, I define the top (bottom) industries as those with the top 5 (bottom 5) industry average returns. The results are reported in Panel D. The results are consistent with those using Fama-French 10 industry classification. Moreover, Column (5) implies that the economic magnitude of the experience effect on Fama-French 48 industry level is also close to that on Fama-French 10 industry level shown in Table 2: the effect of a good experience in Fama-French 48 industry corresponds to an increase of 14.08 percent ($= 0.27 \text{ pp}/1.96 \text{ pp}$) in the probability of future purchases in that industry if normalized by the average probability of buying new stocks in an industry. This evidence may suggest that the effect of good experiences within a category do not become stronger when the category switches from 10 industry to a much finer category - 48 industry. I will further analyze how the experience effect varies with different levels of categorization in Section 4.

3.4 Dynamic Effects

I have so far investigated the influence of prior stock investment experience on a one year horizon. In this section, I will further explore whether the effect lasts for a longer time, and how the significance and magnitude change for lagged experience of different time horizons.

To exhibit more details about the dynamic effects, I switch to use a monthly window in this section. Given using rolling window regression, the larger size of the experience window reduces the sample size. I will investigate the effects of investment experience in the last 36 months and hence start the decision window from January 1994. This choice represents a compromise between having a reasonable length of history of experience and having enough households in the sample. Since some households open the account or stop trading during the sample period, their trading records do not cover the whole sample period and may not be applicable for the test of the long term effects of experience. Therefore, in this section, I restrict the sample to the households who have already opened accounts since 1991 and still actively trade in 1996.

For each household, I use the data of 36 months before the decision month to create the lagged monthly experience related variables and corresponding control variables. The dependent variable indicates the investment decision within a month ($t + 1$) after January 1994, which is a dummy variable equal to one when the household purchase new stocks (never bought in the past 36 months) in one industry during that month. The specification is written as a probit model as follows:

$$P(B_{hI,t+1}^{New} = 1) = \Phi(\beta_0 + \sum_{m=1}^{36} \beta_{1m} Exp_{hI,t+1-m} + \sum_{m=1}^{36} \beta_{2m} Goodexp_{hI,t+1-m} + \Gamma' X) \quad (4)$$

The coefficients on lagged good experience, β_{1m} , are plotted in Figure 3. The figure shows that the effect of experience decreases as the experience happened during earlier periods. The coefficients are significantly positive for the experience in the recent 17 months (except the most recent one), but statistically insignificant when it goes further than 17 months. The effects of monthly experiences within the past year are on average 0.40 pp. This magnitude is lower than that in the baseline model because of the change of window frequency to a

monthly level. However, when normalized by the average probability of buying new stocks in an industry in the decision month 2.36 pp, the magnitude still maintains in the same level that good experience increases the propensity of buying new stocks in the same industry by about 16.95 percent ($= 0.40 \text{ pp}/2.36 \text{ pp}$).

3.5 Summary

Investors' future purchases are influenced by their past experienced returns. If they experience a higher return in an industry, their propensity of purchasing new stocks in that industry increases. This effect is robust after controlling for industry momentum trading, wealth effect, investor heterogeneity and mental accounting. The results remain the same when using different measures of experience, subsample, industry classification. The influence of experienced outcomes becomes weaker when the experiences go further in the past; the experience from 16 months ago or earlier has insignificant effect.

4 Underlying Mechanisms

The results of the baseline model show a significantly positive relationship between good experience and probability of buying new stocks in the same industry. This section further examines the underlying mechanisms: does the relationship vary with the degrees of investor sophistication, portfolio diversification and categorization?

4.1 Investor Sophistication

The influence of experienced outcomes may vary with investor sophistication. If the experience hypothesis is true, the influence of experienced outcomes on future stock purchase decisions is more likely to be stronger for the less sophisticated investors, because given longer history of more sophisticated investors, the recent good or bad experienced outcomes may be canceled by previous ones and have less influence. However, it can also go in the opposite direction if the alternative story of investment ability is true. If the positive relation between past experienced return and future purchases is created by the better performance and more frequent trading of investors with higher investment ability, this relation should be stronger for the investors with higher sophistication.

I use self-reported experience/knowledge in the dataset as a proxy of investor sophistication. When the household opened an account, they would be asked to fill out a form to report their perception of their experience or knowledge in investment. The four levels they could be used to classify themselves are: extensive, high, limited and none knowledge about investing.⁸

I conduct the regressions of specification (3) separately for each group of investors and see how the coefficients change among the subgroups. The results are presented in Table 4. As the table shows, the marginal effect of good experience becomes weaker as investors become more sophisticated. For the most sophisticated subgroup with extensive knowledge about investing, the coefficient is insignificant.⁹

This evidence is more consistent with the experience hypothesis. The most sophisticated investors have a long history of experiences, and may obtain specialized skills in investing certain industries through their historical experienced outcomes. They are more likely to make new purchases in these specialized industries but are not affected much by their recent experience outcomes. In contrast, the least sophisticated ones are new to the stock investment and do not have a long history of experience to cancel out recent experienced outcomes, therefore are more likely to switch across industries depending on whether their experienced outcomes in that industry are good or bad.

4.2 Portfolio Diversification

I next explore the variation in portfolio diversification. Investors with different diversification in their portfolio may take different approach to invest. The investors with more concentrated portfolios may devote themselves to seeking stocks with extraordinary alpha, while some others attempt to diversify their portfolio to reduce idiosyncratic risk. I hypothesize that past investment experience matters more to those alpha seekers relative to diversification oriented investors. People who put all their eggs in one basket have to pick

⁸Because of the missing self-reported data, the sample used in this part is much smaller than the whole sample, but in unreported tables, I show that the main summary statistics of distribution of trades among industries and distribution of experience across industries and years don't change much.

⁹I also run one regression including the observations of all the groups, and interact the sophistication group with past experience. The results show that the difference of the marginal effect of good experience is statistically significant between the none knowledge group and the extensive knowledge group.

the best basket and their past experience could play a role in their decision making. In contrast, people who put their eggs in multiple baskets would care less about which baskets they choose. Therefore, I expect to see that the relationship between past experienced outcome and future new purchases in the same industry is more significant for the investors with more concentrated portfolios.

I test this hypothesis by adding an interaction term between experience-related variable and a diversification measure into the baseline model, the specification is as follows:

$$\begin{aligned}
 P(B_{hI,t+1}^{New} = 1) = & \Phi(\beta_0 + \beta_1 Exp_{hIt} + \beta_2 Goodexp_{hIt} + \beta_3 Exp_{hIt} \cdot Diversified_{ht} \\
 & + \beta_4 Goodexp_{hIt} \cdot Diversified_{ht} + \beta_5 Diversified_{ht} + \Gamma' X_{hIt}) \quad (5)
 \end{aligned}$$

I use two measures for portfolio diversification following [Ivkovic, Sialm, and Weisbenner \(2008\)](#). The first measure is a portfolio Herfindahl Index (HI), which is defined as the sum of the squared weights of each stock k in the household stock portfolio (w_k), i.e. $HI = \sum_k w_k^2$. The more diversified the portfolio is, the Herfindahl index is smaller. If a household owns only one common stock, the Herfindahl Index reaches its maximum and equals to one. Therefore, I define the dummy variable $Diversified_{it}$ equal to one when the Herfindahl index of investor i 's portfolio is less than the median in the sample. The second measure is the number of stocks held in the portfolio. A portfolio with more stocks is considered as more diversified. Similarly, I define the $Diversified_{it}$ variable as equal to one when the number of stocks held is greater than 5, which is the median of the sample.

According to results in [Table 5](#), the influence of experienced outcomes is stronger among the investors who hold more concentrated portfolios, since the coefficient on the interaction term between good experience and diversified portfolio is negative and statistically significant. As for investors with concentrated portfolios, good experience in an industry increases their probability of buying new stocks in the same industry by 1.32 pp. This effect is dampened by about 2/3 when the portfolio is more diversified.

4.3 Categorization

Up to this point, I have been using Fama-French 10 industry as the classification of industries. To shed more light on to what extent investors categorize stocks, I will explore how the impact of past experiences varies if I apply a broader categorization or a finer categorization.

4.3.A Spill-over Effect across Industry

I start by testing whether past experiences of Fama-French 10 industries have spill-over effects on the purchases of other industries. The light blue and white bars in Figure 2 display the probability of purchasing new stocks in other industries by sorting households' experienced market-adjusted returns in the past year. The right-most bars (group 5) correspond to the top 20% experienced returns, while the left-most bars (group 1) correspond to the bottom 20% experienced returns. I create a distance measure among industries to select, among other industries, the one most similar to the experienced industry and the one most different from the experienced industry. For each year, the distance is calculated by averaging daily absolute difference between the stock returns of two industries. Suppose there are N trading days in one year, the distance between industry I and industry J (both defined by Fama-French 10-industry) is described as $D_{IJ} = \frac{1}{N} \sum_{t=1}^N |R_{It} - R_{Jt}|$.¹⁰ The matches are intuitive as well. For example, for most of the years, the most similar industry to the industry of oil, gas, and coal extraction and products is the utility industry, while the most different industry is the high technology industry.

Figure 2(a) plots the original probability of purchasing new stocks. The probability in the most similar industry slightly increases as the experienced return goes up, while the probability in the most different industry almost maintains itself at the same level across different levels of experienced returns. In a word, the effect of past experienced returns becomes increasingly weaker when the industry becomes more different from the experienced industry. Using the residuals from regressing the purchase decision on the controls, Figure 2(b) confirms this evidence.

¹⁰I also experimente with another distance measure $D_{IJ} = \frac{1}{N} \sqrt{\sum_{t=1}^N (R_{It} - R_{Jt})^2}$, and the results remain the same.

Table 6 uses a regression framework to test whether spill-over effect on other industries is significant while controlling for the wealth effect, investor heterogeneity, year and industry effects. As in Figure 2, I consider the effect on two other industries: the most similar to the experienced industry and the most different. According to Table 6, compared to the effect of good experience on the purchases in the same industry, the effect in the most similar industry is only marginally significant, and the magnitude is far smaller than that in the same industry; the effect in the most different industry is even smaller and insignificant. If we also take a look at all the control variables, the significance and magnitude do not change much no matter if it is for the experienced industry or for other industries. This result is consistent with our intuition that wealth effect and investor characteristics should have the same effect on each industry.

The evidence in Figure 2 and Table 6 suggests that the impact of past experience does not spill over, or if it does only in a tiny amount, to other industries.

4.3.B Categorical Thinking within Industry

In this section, I further investigate whether investors categorize stocks more finely than Fama-French 10 Industry. Specifically, I will consider finer categorization in two dimensions. In the first dimension, the investor may have a finer categorization through the representativeness of the stocks in which the investor has experience. If investors purchased stocks that are representative of the industry, is the impact of experienced outcomes larger compared to if investors invested in unrepresentative stocks? I construct a measure for the representativeness of each stock. The measure for an experienced stock j in industry I is computed as the correlation between the stock return r_{jt} and the equal-weighted industry return R_{It} .¹¹ A higher correlation indicates that the experienced stock is more representative of the stocks in the same industry. The representative stocks identified by this measure are intuitive as well. For example, the representative stocks in the industry of wholesale, retail and some services include May Department Store, Target, Home Depot, etc., while the non-representative stocks in this industry include Perfumania, Skyline Chili Inc., etc.

I divide the experiences into two groups according to the representativeness of the ex-

¹¹I use equal-weighted industry return to address the concern that the measure would be biased toward large firms if I use value-weighted industry return.

perienced stocks. After removing the effects of controls, Figure 4 displays the probability of purchasing new stocks for experiences on representative stocks and non-representative stocks, respectively. According to the figure, the probability of purchasing new stocks increases with experienced returns for both groups, and the probability increases by a larger amount for the more representative group. It means that investors are influenced more by experiencing more representative stocks. If they profit from investing in Target, they may have higher probability to buy other stocks in the wholesale, retail and services industry again. In contrast, a good experience from investing in Perfumania may not have as much impact.

The representativeness captures one dimension of finer categorization based on the statistical feature of stock returns, and individual investors may perceive finer categorization in a more intuitive way. They may further divide stocks into subcategories beyond Fama-French 10-industry classification, such as Fama-French 48-industry. If the investors consider stocks under Fama-French 48-industry classification, it should be true that the influence of past experience within the same subindustry (Fama-French 48 Industry) is stronger than that on another subindustry, even though these two subindustries belong to the same Fama-French 10 industry.

The industries under Fama-French 48-industry classification may correspond to multiple industries under Fama-French 10-industry classification. To keep the test clean, I only consider one specific industry by Fama-French 10-industry classification in the test. Table 7 reports the results using the data from the Manufacturing industry according to Fama-French 10-industry classification. As Column (1) shows, good experience in a subindustry is significantly positive related to subsequent purchases of new stocks in the same subindustry. Column (2) presents the impact on another random subindustry in the Manufacturing industry. The coefficient on the good experience variable is also significantly positive and very close to the impact on the experienced subindustry (Column (1)). If an investor gains a good experience in one subindustry, she may buy new stocks in another subindustry within the same industry with the same probability as she may buy in the experienced subindustry. This evidence may suggest that investors do not think as finer as Fama-French 48-industry classification. To complement the test, I also report the impact on a random subindustry

outside Manufacturing (Column (3)). The coefficient is insignificant, which is consistent with Table 6: the impact does not spill over beyond Fama-French 10 industry.

Overall, I find that, within Fama-French 10 industry, experiences on more representative stocks have a stronger impact on future purchases in the same industry; but the impact does not seem to become stronger if we further segment Fama-French 10 industry to Fama-French 48 industry.

4.4 Discussion

This experience effect is consistent with two possible mechanisms.¹² The first mechanism is that investors may put more weight on their experience outcomes than on other available historical information when undertaking a Bayesian updating their beliefs about stocks in the same industry. This is a natural explanation for why experience outcomes affect purchase decisions. Good experiences drive posterior expected returns upward, and consequently investors are more likely to buy new stocks in that industry. Another explanation is that investors may construe their experienced outcomes as indications of their ability to invest in a particular industry. Therefore, an investor may (correctly or falsely) learn that she has an advantage in investment in the energy industry when her energy stocks outperform the market; she then buys more stocks in the energy industry to take advantage of this perceived ability. Table 8 provides suggestive evidence for this mechanism. I run a Tobit regression of the number of trades in one industry (including both buys and sales) on the previous year's experience-related variables pertinent to that industry, controlling for industry average and individual characteristics. The results show that investors will trade more frequently in an industry after they have had good experience in that industry. This evidence is consistent with the hypothesis that investors may perceive that they are better at picking stocks in the industries in which they had good experiences.

These two mechanisms are not mutually exclusive, and both of them could have implications on investors' welfare. Under the first mechanism, investors' past experiences may bias their belief and lead them to miss good investment opportunities. As one practitioner

¹²Experience could also influence investors' behavior through other channels. For example, in the context of Barber et al. (2010), experienced buying or selling price could affect investors' reference point and further influence investors' action since investors tend to avoid anticipated regret. Because my research focuses on the decision of buying stocks not previously owned, the role of affecting reference point may not apply.

notes, “the problem is that in accumulating experience, he also acquires prejudices against industries and stocks because he has lost money in them. It is easy to ... become an investment bigot with a closed mind on many subjects.”¹³ Does this behavior hurt investors’ performances? Figure 5 compares the returns of actual purchases by the investors with the returns of two hypothetical strategies: (1) the index of industries in which the investors have picked stocks; and (2) the index of industries which investors could pick but choose not to. I form calendar-time portfolios corresponding to these strategies. For each day, the calendar-time portfolio of actual purchases is constructed to include all stocks bought by households within the prior 21 trading days; and the portfolio of (not) invested industry is formed by investing in indices of the industries in which households have (not) bought stocks during the prior 21 trading days.¹⁴ The weight is equally allocated to each stock-household (industry-household) pair. Figure 5a displays the monthly abnormal returns of these three portfolios controlling for Fama-French three factors (Fama and French, 1993). According to the figure, past experiences do not seem matter much for each portfolio, and in general returns of actual purchases are negative, while the returns of industry averages are positive. For comparison, Figure 5b plots the difference of the monthly abnormal returns between actual purchases and the industries invested by households. As the figure shows, the stocks picked by investors earn a much lower return compared to the average returns of corresponding industries. The industry-adjusted returns of picked stocks are statistically significant and negative, indicating that on average individual investors have inferior stock-picking ability. Additionally, Figure 5c plots the difference of the monthly abnormal returns between industries that are invested and not invested by households. According to the figure, these two returns are quite close to each other and the difference between the two is statistically insignificant regardless of whether investors bought in the industries with good, bad, or no experiences. In other words, even though investors tend to purchase more in industries regarding which they have had good experience, they are not missing much as a result of bias towards good experience industries. Individual investors would not systematically pick

¹³See <http://dailyreckoning.com/the-worst-possible-investing-mistake/>

¹⁴I exclude the stocks bought on the day of forming the portfolio. It is to address the concern that if investors tend to buy stocks after observing good returns of the industry average of that day, then including the stocks bought on the day of forming the portfolio may mechanically observe the returns of actual purchases lower than the industry average returns by the way constructing the portfolio.

the wrong industries for investment. What may actually hurt their performance is that they appear to systematically pick the wrong stocks for investment.

Considering the evidence that on average individual investors appear to have inferior stock-picking ability as shown in Figure 5, it is more likely that investors incorrectly assess their ability as high because of experienced positive outcomes. As a result, investors are more likely to trade in industries in which they have had good experiences as well as to increase their trading frequency after good experiences. Their performance will then decline, both as a consequence of their inept stock picking ability and the increased transaction costs of their more frequent trades. Odean and Gervais (2001) point out that investors become overconfident because they tend to overestimate their trading ability based on their successes and downplay their failures. The evidence in this paper raises another mechanism of building up overconfidence. Investors categorize their portfolio and infer their ability to trade in each category from their past performances. This categorization increases the chances that they will find good performances among categories they invested in and mistakenly attribute their success to talent for investing in those categories. This, in turn, causes the inertia which decreases the likelihood of giving up a stock picking strategy before diversifying their portfolio or exiting stock market.

5 Conclusion

This paper investigates the influence of personal investment experience on subsequent stock selection decisions. Using trading records data for households from a large discount broker collected over the period between 1991 and 1996, I demonstrate that investors have a higher propensity to purchase new stocks in an industry if during the past in that same industry they earned positive excess returns. I also provide evidence that the significance and magnitude of the influence of prior experienced outcomes varies over different time horizons, categorizations, degrees of investors' sophistication and investors' portfolio diversification.

The influence of past investment experience may result from investors overweighing their own experience relative to other available historical information when updating their beliefs about stocks in a category (e.g. the industry examined in this paper). The results may also be explained by investors learning about their ability of picking stocks in a certain category

from their experiences. I provide evidence that investors may not systematically miss good opportunities as a consequence of this biased belief updating. Nevertheless, the good experience by inept investors in a given category may lead them to assume that they possess insight regarding decisions in this category. In this case, finer categorization will delay the exit of such inferior investors and cause welfare loss. Furthermore, the aggregation of the influences of investors' personal investment experiences may have a systematic effect on asset pricing; for example, this experience effect may provide a source of industry momentum effect that is different from the existing explanations in the literature. Finally, if past experience outcomes influence investors' decisions by affecting information acquisition or the constitution of their information set, we may be able to make richer predictions about the information incorporation in the market.

References

- Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55, 773–806.
- Barber, B. M. and T. Odean (2007). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21(2), 785–818.
- Barber, B. M., T. Odean, and M. Strahilevitz (2010). Once burned, twice shy: How pride and regret affect the repurchase of stocks previously sold. *Working Paper*.
- Barber, B. M., T. Odean, and N. Zhu (2009). Systematic noise. *Journal of Financial Markets* 12(4), 547–569.
- Barberis, N. and A. Shleifer (2003). Style investing. *Journal of Financial Economics* 68, 161–199.
- Campbell, J. Y., T. Ramadorai, and B. Ranish (2013). Getting better: Learning to invest in an emerging stock market. *Working Paper*.
- Chan, L. K., H. L. Chen, and J. Lakonishok (2002). On mutual fund investment styles. *The Review of Financial Studies* 15(5), 1407–1437.
- Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2009). Reinforcement learning and savings behavior. *Journal of Finance* 64(6), 2515–2534.
- DeLong, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). Positive feedback investment strategies and destabilising rational speculation. *Journal of Finance* 45, 375–395.
- Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Fama, E. F. and K. R. French (1997). Industry costs of equity. *Journal of Financial Economics* 43, 153–193.

- French, K. R. and J. M. Poterba (1991). International diversification and international equity markets. *American Economic Review* 81(2), 222–226.
- Gallagher, J. (2012). Learning about an infrequent event: Evidence from flood insurance take-up in the us. *Working Paper*.
- Graham, J. R. and K. Narasimhan (2004). Corporate survival and managerial experiences during the great depression. *Duke University Working Paper*.
- Hertwig, R., G. Barron, E. U. Weber, and I. Erev (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science* 15, 534–539.
- Hong, H. and J. C. Stein (1999). A unified theory of underreaction, momentum trading and overreaction in asset markets. *Journal of Finance* 54, 2143–2184.
- Ivkovic, Z., C. Sialm, and S. Weisbenner (2008). Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis* 43, 613–656.
- Jackson, A. (2003). The aggregate behavior of individual investors. *Working Paper*.
- Kaustia, M. and S. Knupfer (2008). Do investors overweight personal experience? evidence from ipo subscriptions. *Journal of Finance* 63, 2679–2702.
- Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics* 126(1), 373–416.
- Malmendier, U. and G. Tate (2005). Ceo overconfidence and corporate investment. *Journal of Finance* 60, 2661–2700.
- Malmendier, U., G. Tate, and J. Yan (2011). Overconfidence and early-life experiences: The impact of managerial traits on corporate financial policies. *Journal of Finance*, Forthcoming.
- Moskowitz, T. J. and M. Grinblatt (1999). Do industries explain momentum? *Journal of Finance* 4, 1249–1290.

- Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance* 53, 1775–1798.
- Odean, T. (1999). Do investors trade too much? *American Economic Review* 89(5), 1279–1298.
- Odean, T. and S. Gervais (2001). Learning to be overconfident. *Review of Financial Studies* 14, 1–27.
- Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics* 80, 563–602.
- Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making* 12, 183–206.
- Weber, E. U., U. Bockenholt, D. Hilton, and B. Wallace (1993). Determinants of diagnostic hypothesis generation: Effects of information, base rates, and experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19, 1151–1164.

Figure 1: Motivating Example: Difference in Stock Holdings by Industry of Young and Old Groups vs Difference in Their Experienced Industry Returns

The stock holdings of an age group in industry i are measured by the fraction of households investing in industry i in a certain year within the age group. The young group is defined as the households with age between 30 and 40, while the old group is defined as the households with age between 50 and 60. The vertical axis denotes the difference in the stock holdings by industry of these two groups. The horizontal axis denotes the difference in their experienced industry returns. The experience of the young group in industry i is measured as the average of industry i returns over the prior 10 years, while the experience of the old group is measured over the prior 30 years. Each observation corresponds to a year-industry pair. Industry is classified according to Fama-French 10 industry classification. “High Tech” industry is excluded in this analysis.

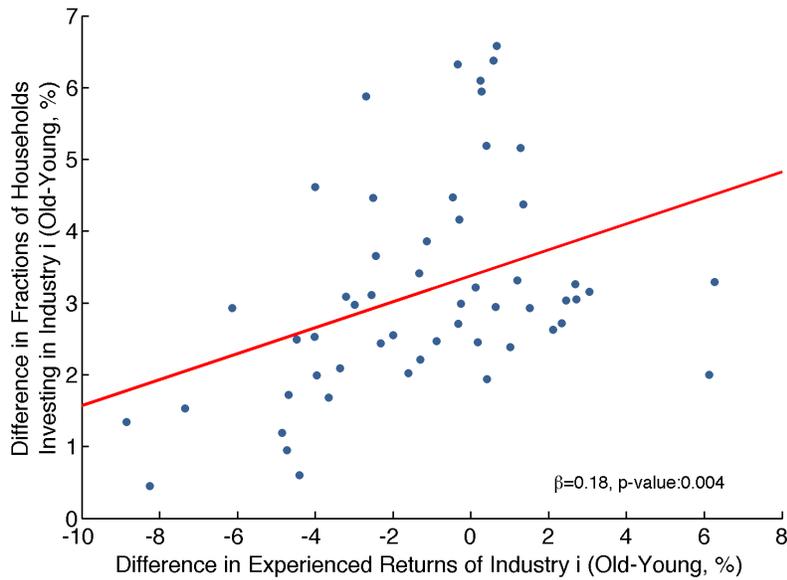
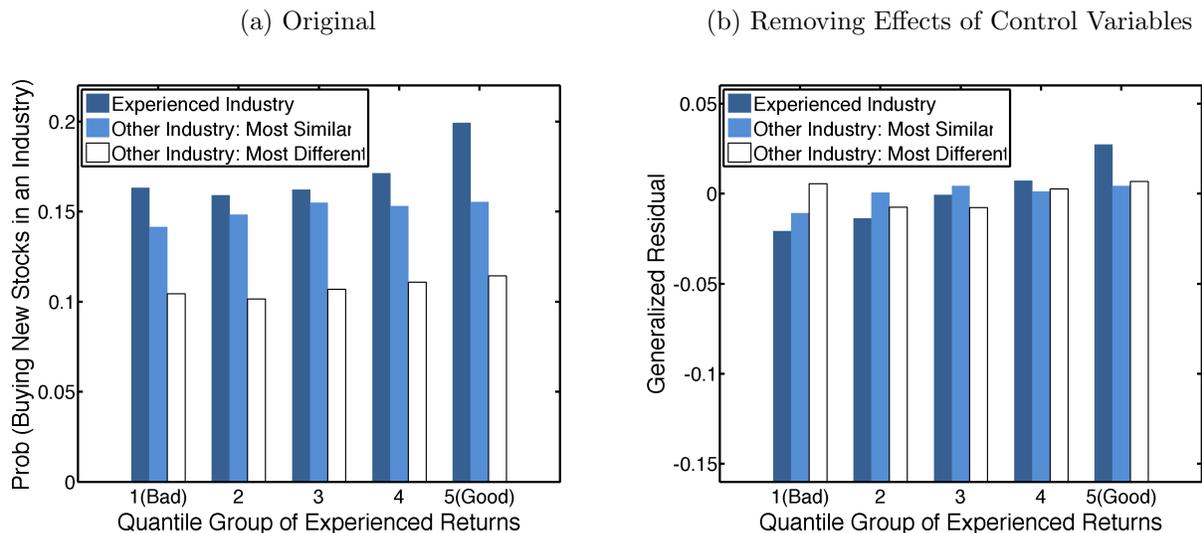


Figure 2: The Impact of Personal Investment Experience in an Industry on Subsequent Investment in Experienced Industry vs Other Industry

The observations are sorted by the value weighted average annualized excess return on the investment in an industry and divided into five groups. Group 1 has the lowest experienced return while group 5 has the highest. Figure (a) plots the original probability of buying new stocks, i.e. the percent of households buying new stocks in an industry. Figure (b) plots the generalized residuals from a probit model of regressing purchasing new stocks in an industry on controls of the baseline model.¹⁵The dark blue bars correspond to the experienced industry. The light blue and white bars correspond to the industry that is the most similar to or the most different from the experienced industry, respectively. The most similar and the most different industries are selected by measuring the distance between the stock returns in that industry and those in the experienced industry. The distance is measured by averaging daily absolute difference between the stock returns of two industries.



¹⁵The generalized residual of the probit model $Pr(Y_i = 1) = \Phi(X_i'\beta)$ is computed as:

$$\frac{Y_i - \Phi(X_i'\beta)}{\Phi(X_i'\beta)(1 - \Phi(X_i'\beta))} \phi(X_i'\beta)$$

where $\phi(\cdot)$ is the density function of the normal distribution and $\Phi(\cdot)$ is the cumulative distribution.

Figure 3: Dynamic Effects of Monthly Lagged Industry Experienced Outcomes

This figure shows the dynamic effects of experienced outcomes in the past 36 months on subsequent purchases of new stocks. The sample includes all the investors who already have investment from 1991 and still trade in 1996 in the dataset. The solid line connects the estimated coefficients of a rolling-window probit regression of purchasing new stocks on monthly experienced outcomes during the past 36 months. Time 0 corresponds to the decision window during which the decisions whether purchasing new stocks are made. The decision window starts from January 1994 through December 1996. The regressions are controlled for industry momentum. The broken lines show the 5% and 95% confidence intervals.

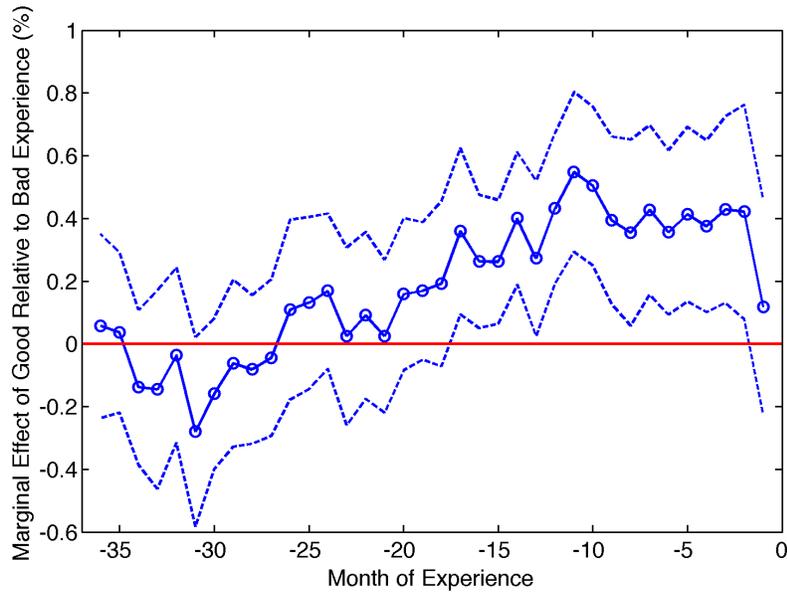
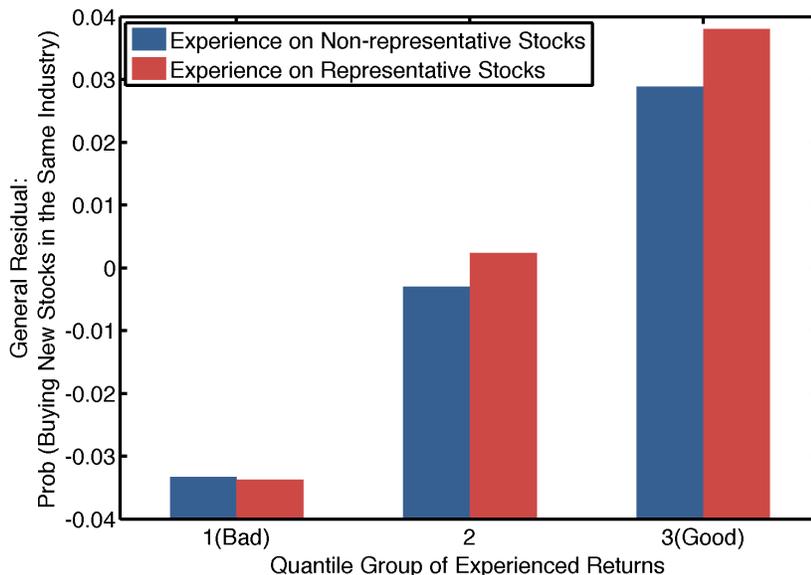


Figure 4: The Impact of Investment Experience: Variation in Representativeness of Experienced Stocks

The observations are sorted by the value weighted average annualized excess return on the investment in an industry and divided into three groups. Group 1 has the lowest experienced return while group 3 has the highest. The figure plots the generalized residuals from a probit model of regressing purchasing new stocks in an industry on controls of the baseline model.¹⁶ The blue bar corresponds to the observations with experiences in the non-representative stocks, while the red bar corresponds to those with experiences in the representative stocks. The sample only includes the observations whose representativeness measures are available. The representativeness measure for one experienced stock j in industry I is computed as the correlation between the stock return r_{jt} and the equal-weighted industry return R_{It} . The representative stocks are defined as stocks with representativeness measures greater than the median.



¹⁶The generalized residual of the probit model $Pr(Y_i = 1) = \Phi(X_i'\beta)$ is computed as:

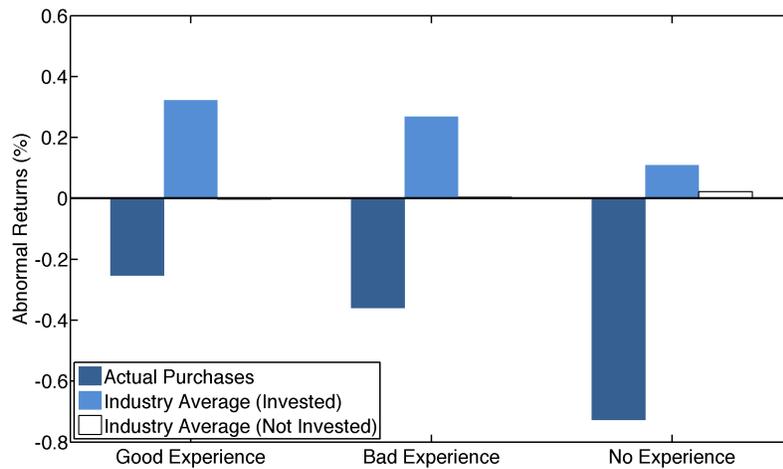
$$\frac{Y_i - \Phi(X_i'\beta)}{\Phi(X_i'\beta)(1 - \Phi(X_i'\beta))} \phi(X_i'\beta)$$

where $\phi(\cdot)$ is the density function of the normal distribution and $\Phi(\cdot)$ is the cumulative distribution.

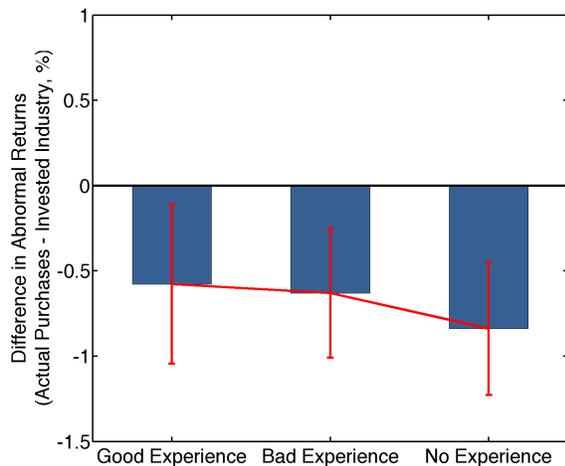
Figure 5: Performance Analysis (Calendar Time Portfolio)

Figure (a) plots the monthly abnormal returns of calendar time portfolios controlling for Fama-French three factors (Fama and French, 1993). At each day, the calendar-time portfolio of actual purchases is formed to include all stocks bought by households within prior 21 trading days; the portfolio of (not) invested industry is formed to include all industries within which households have (not) bought stocks during prior 21 trading days. The returns of the portfolio are then accumulated within each month to obtain monthly returns. Figure (b) plots the monthly abnormal returns of going long the calendar time portfolio of actual purchases and short the portfolio of invested industries; Figure (c) plots the monthly abnormal returns of going long the portfolio of invested industries and short the portfolio of not invested industries.

(a) Portfolio Performances



(b) Actual Purchases - Invested Industry



(c) Invested Industry - Not Invested Industry

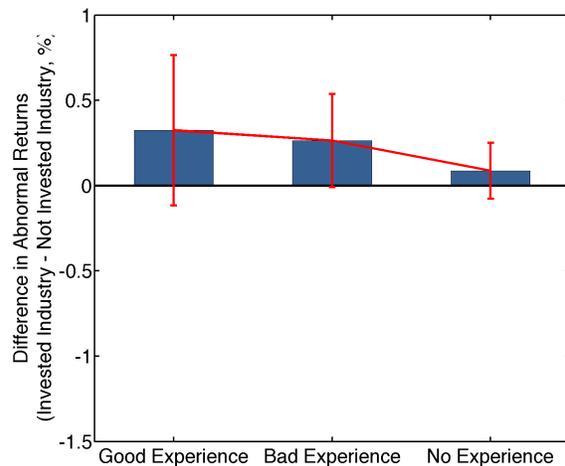


Table 1: Summary Statistics

Panel A reports the number and percentage of trades purchasing other stocks not previously owned and trades repurchasing the stocks previously owned. The sample period includes all the years which are used as decision windows in the baseline model (1992-1996). Panel B reports the percentage of households investing in each industry groupings from 1991-1995 for the full sample. Panel C reports the percentage of household's good(bad) experience in each industry groupings from 1991-1995 for the full sample. Good(bad) Experience is measured as the experienced market-adjusted return in the industry is greater than(smaller than or equal to) zero. The sample period includes all the years which are used as experience windows in the baseline model.

Panel A: Distribution of Household Stock Purchases, 1992-1996

	1992	1993	1994	1995	1996	Total
Purchase Other Stocks	77,057	70,912	57,626	68,519	72,727	346,841
Repurchase Previously Owned Stocks	10,455	12,013	12,200	10,946	13,839	59,453
Total	87,512	82,925	69,826	79,465	86,566	406,294
Purchase Other Stocks	88.1%	85.5%	82.5%	86.2%	84.0%	85.4%
Repurchase Previously Owned Stocks	11.9%	14.5%	17.5%	13.8%	16.0%	14.6%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Panel B: Distribution of Household Stock Investment Experience Across Industry, 1991-1995

	1991	1992	1993	1994	1995	Total
(1) Consumer nondurables (%)	8.7	8.7	9.5	9.3	8.8	9.0
(2) Consumer durables (%)	4.6	5.0	4.4	4.6	4.9	4.7
(3) Manufacturing (%)	11.9	11.0	10.5	10.1	10.6	10.7
(4) Oil, gas, and coal extraction and products (%)	4.5	4.8	4.5	4.2	4.2	4.4
(5) High technology (%)	17.4	17.1	16.6	16.3	17.0	16.8
(6) Telephone and television transmission (%)	4.8	5.0	5.6	6.5	6.9	5.9
(7) Wholesale, retail, and some services (%)	10.3	10.7	11.3	11.7	12.0	11.3
(8) Health care, medical equipment, and drugs (%)	12.5	14.7	15.3	14.4	13.3	14.1
(9) Utilities (%)	5.1	5.8	5.8	6.4	6.1	5.9
(10) Others (%)	20.2	17.2	16.6	16.5	16.2	17.0
N	58,392	86,513	101,525	107,534	116,928	470,892

Panel C: Household Stock Investment Outcomes, 1991-1995

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Total
	Consumer non- durables	Consumer durables	Manufacturing	Oil, gas, and coal extraction and products	High technology	Telephone and television transmis- sion	Wholesale, retail, and some services	Health care, medical equipment, and drugs	Utilities	Others	
<i>Year = 1991</i>											
Bad (< 0) (%)	50.8	65.3	57.6	81.2	60.6	52.9	44.8	40.2	44.0	61.2	55.3
Good (> 0) (%)	49.2	34.7	42.4	18.8	39.4	47.1	55.2	59.8	56.0	38.8	44.7
N	5,058	2,715	6,927	2,633	10,177	2,793	6,012	7,297	2,959	11,821	58,392
<i>Year = 1992</i>											
Bad (< 0) (%)	55.1	34.9	53.3	63.9	53.6	33.7	45.2	74.4	57.2	43.4	52.9
Good (> 0) (%)	44.9	65.1	46.7	36.1	46.4	66.3	54.8	25.6	42.8	56.6	47.1
N	7,567	4,334	9,510	4,188	14,773	4,312	9,214	12,725	5,035	14,855	86,513
<i>Year = 1993</i>											
Bad (< 0) (%)	68.8	28.7	44.2	41.4	45.6	47.1	60.1	74.8	58.3	38.4	52.5
Good (> 0) (%)	31.2	71.3	55.8	58.6	54.4	52.9	39.9	25.2	41.7	61.6	47.5
N	9,633	4,429	10,621	4,579	16,871	5,691	11,447	15,575	5,870	16,809	101,525
<i>Year = 1994</i>											
Bad (< 0) (%)	53.1	67.3	45.9	49.3	37.9	69.2	64.2	39.7	78.6	62.4	54.0
Good (> 0) (%)	46.9	32.7	54.1	50.7	62.1	30.8	35.8	60.3	21.4	37.6	46.0
N	9,981	4,910	10,908	4,537	17,558	6,960	12,600	15,438	6,893	17,749	107,534
<i>Year = 1995</i>											
Bad (< 0) (%)	50.4	80.4	61.9	64.9	62.6	49.7	70.2	34.0	80.5	53.3	58.2
Good (> 0) (%)	49.6	19.6	38.1	35.1	37.4	50.3	29.8	66.0	19.5	46.7	41.8
N	10,302	5,706	12,415	4,961	19,928	8,019	14,038	15,547	7,111	18,901	116,928

Table 2: Investment Experience and the Propensity to Purchase New Stocks in the Same Industry (Fama-French 10-industry Classification)

This table reports maximum likelihood regression results for probit regressions. The results are reported as marginal effects of independent variables. Each observation corresponds to one household and one industry, regardless whether the household previously owned stocks in the industry or not. The dependent variable is based on a dummy variable coded one when a household purchases stocks not previously owned in the industry. “Experience dummy” is coded one if the household owns stocks in the industry in the past year. Other independent variables are dummy variables related to personal experience, industry average, portfolio size and individual characteristics variables. Both personal experience variables and industry average variables are based on the value of market-adjusted (experienced or industry average) return. “Good experience”(“Good industry”) is coded one when the market-adjusted experienced (industry average) return is greater than 0. “Top experience”(“Bottom experience”) is coded one when the market-adjusted experienced return is above (below) the 90th percentile (10th percentile) of the sample. “Top 1 industry”(“Bottom 1 industry”) is coded one when the industry average return is the highest(lowest) among the 10 industries. “Increase of portfolio size” equals to one if the size of household’s portfolio increases in the past year. The three individual characteristics variables are created from the beginning-of-month position data. They denote the average monthly number of stocks, the average monthly size and the average monthly turnover rate in the past year. Standard errors, shown in parentheses, are clustered by industry-year level. *10%, **5%, ***1% significance.

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Buy New Stocks (not Owned in the Past Year) in One Industry				
Experience Dummy	0.07982*** (0.00559)	0.08760*** (0.00640)	0.09239*** (0.00810)	0.06858*** (0.00943)	0.06571*** (0.00937)
Good Experience (> 0)	0.01148*** (0.00266)	0.01069*** (0.00246)	0.01271*** (0.00283)	0.01260*** (0.00298)	0.01856*** (0.00312)
Top Experience (over the 90th percentile)	0.01925*** (0.00492)	0.01879*** (0.00488)	0.02011*** (0.00472)	0.01459*** (0.00415)	
Bottom Experience (below the 10th percentile)	-0.01288*** (0.00211)	-0.01263*** (0.00204)	-0.01468*** (0.00251)	-0.01185*** (0.00333)	
Good Industry (> 0)	0.00396 (0.00580)	0.00575 (0.00444)	0.00653 (0.00525)	0.00621 (0.00688)	0.00618 (0.00689)
Top 1 Industry		0.02282** (0.01128)	0.02914** (0.01276)	0.03682** (0.01433)	0.03703** (0.01446)
Bottom 1 Industry		0.02667** (0.01164)	0.02845** (0.01312)	0.03109** (0.01490)	0.03132** (0.01499)
Increase of Portfolio Size			0.01833*** (0.00202)	0.01489*** (0.00168)	0.01482*** (0.00168)
Average Num of Stocks > 5				0.05680*** (0.00204)	0.05685*** (0.00205)
log(Average Portfolio Size)				0.01528*** (0.00071)	0.01539*** (0.00072)
log(Average Turnover Rate)				0.00832*** (0.00076)	0.00851*** (0.00078)
Industry Effect	Yes	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes	Yes
Avg Prob (Buying New Stocks in an Industry)	0.0846	0.0846	0.0930	0.1118	0.1118
Observations	1,550,980	1,550,980	1,094,860	777,630	777,630

Table 3: Robustness Tests

This table reports maximum likelihood regression results for probit regressions. The results are reported as marginal effects of independent variables. Panel A exploits the level of market-adjusted experienced and industry average returns as the experience variable and industry average variables, instead of dummy variables. Panel B only includes the observations with experiences in the industry in the past year. In Panel C, the dependent variable is defined as one if the household purchases stocks not previously owned in the industry and does not have sales in that industry within 30 days before the purchases. Panel D creates investment experiences based on Fama-French 48 industry classification. Without special specification, each observation corresponds to a household/industry/year pair, regardless whether the household previously owned stocks in the industry or not. The dependent variable is based on a dummy variable coded one when a household purchases stocks not previously owned in the industry. “Experience dummy” is coded one if the household owns stocks in the industry in the past year. Other independent variables are dummy variables related to personal experience, industry average, portfolio size and individual characteristics variables. Both personal experience variables and industry average variables are based on the value of market-adjusted (experienced or industry average) return. “Good experience”(“Good industry”) is coded one when the market-adjusted experienced (industry average) return is greater than 0. “Top experience”(“Bottom experience”) is coded one when the market-adjusted experienced return is above (below) the 90th percentile (10th percentile) of the sample. “Top 1 industry”(“Bottom 1 industry”) is coded one when the industry average return is the highest(lowest) among the 10 industries. The wealth effect and individual characteristics controls are defined as in Table 2. Standard errors, shown in parentheses, are clustered by industry-year level. *10%, **5%, ***1% significance.

Panel A: Measure Experienced Outcomes by the Level of Market-adjusted Returns

	(1)	(2)	(3)	(4)
Dependent Variable:	Buy New Stocks (not Owned in the Past Year) in One Industry			
Expeirence Dummy	0.08553*** (0.00356)	0.08912*** (0.00453)	0.09281*** (0.00474)	0.07058*** (0.00495)
Experienced Excess Return	0.00119*** (0.00030)	0.00119*** (0.00029)	0.00121*** (0.00040)	0.00124*** (0.00042)
Industry Average Excess Return	0.02065 (0.02785)	-0.01171 (0.02553)	-0.00949 (0.02769)	-0.00855 (0.03034)
Square of Industry Average Excess Return		0.48734*** (0.16751)	0.50777*** (0.18348)	0.57314*** (0.21884)
Wealth Effect Control	No	No	Yes	Yes
Individual Characteristics Control	No	No	No	Yes
Industry Effect	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes
Observations	1,550,980	1,550,980	1,094,860	777,630

Panel B: Only Industry-Year Observations with Past Experience

	(1)	(2)	(3)	(4)
Dependent Variable:	Buy New Stocks (not Owned in the Past Year) in One Industry			
Good Experience (> 0)	0.01385*** (0.00449)	0.01256*** (0.00437)	0.01494*** (0.00470)	0.01412*** (0.00458)
Top Experience (over the 90th percentile)	0.01125* (0.00613)	0.01006* (0.00598)	0.00979* (0.00578)	0.00375 (0.00415)
Bottom Experience (below the 10th percentile)	-0.01182** (0.00500)	-0.01156** (0.00494)	-0.01486*** (0.00425)	-0.01350*** (0.00451)
Good Industry (> 0)	0.00437 (0.01134)	0.00207 (0.00955)	0.00045 (0.01019)	-0.00068 (0.01083)
Top 1 Industry		0.05638*** (0.02056)	0.07273*** (0.02073)	0.08898*** (0.02073)
Bottom 1 Industry		0.03172 (0.02477)	0.03165 (0.02615)	0.03378 (0.02684)
Wealth Effect Control	No	No	Yes	Yes
Individual Characteristics Control	No	No	No	Yes
Industry Effect	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes
Observations	385,552	385,552	278,427	225,142

Panel C: Robustness Check for Mental Accounting

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Buy New Stocks in One Industry without Recent Sales				
Experience Dummy	0.07781*** (0.006)	0.08581*** (0.006)	0.09048*** (0.008)	0.06635*** (0.009)	0.06466*** (0.009)
Good Experience (> 0)	0.00562** (0.003)	0.00478* (0.003)	0.00669** (0.003)	0.00687** (0.003)	0.00917*** (0.003)
Top Experience (over the 90th percentile)	0.00642* (0.004)	0.00580 (0.004)	0.00536 (0.003)	0.00306 (0.003)	
Bottom Experience (below the 10th percentile)	-0.00683** (0.003)	-0.00658** (0.003)	-0.00834*** (0.003)	-0.00783*** (0.003)	
Good Industry (> 0)	0.00405 (0.006)	0.00574 (0.004)	0.00654 (0.005)	0.00625 (0.007)	0.00628 (0.007)
Top 1 Industry		0.02289** (0.011)	0.02922** (0.013)	0.03681** (0.014)	0.03691** (0.014)
Bottom 1 Industry		0.02660** (0.012)	0.02850** (0.013)	0.03106** (0.015)	0.03118** (0.015)
Wealth Effect Control	No	No	Yes	Yes	Yes
Individual Characteristics Control	No	No	No	Yes	Yes
Industry Effect	Yes	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes	Yes
Avg Prob (Buying New Stocks in an Industry)	0.0831	0.0831	0.0913	0.0808	0.0808
Observations	1,550,980	1,550,980	1,094,860	777,630	777,630

Panel D: Fama-French 48 Industry Classification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Buy New Stocks (not Owned in the Past Year) in One Industry				
Experience Dummy	0.02205*** (0.00217)	0.02349*** (0.00232)	0.02404*** (0.00297)	0.01583*** (0.00252)	0.01550*** (0.00251)
Good Experience (> 0)	0.00163** (0.00072)	0.00154** (0.00061)	0.00190*** (0.00073)	0.00175** (0.00076)	0.00276*** (0.00077)
Top Experience (over the 90th percentile)	0.00385*** (0.00080)	0.00370*** (0.00073)	0.00408*** (0.00078)	0.00304*** (0.00068)	
Bottom Experience (below the 10th percentile)	-0.00120** (0.00054)	-0.00121** (0.00051)	-0.00142** (0.00062)	-0.00121 (0.00075)	
Good Industry (> 0)	0.00159* (0.00096)	0.00095 (0.00116)	0.00143 (0.00141)	0.00189 (0.00160)	0.00189 (0.00160)
Top 5 Industries		0.00313** (0.00125)	0.00341*** (0.00132)	0.00400*** (0.00147)	0.00400*** (0.00147)
Bottom 5 Industries		0.00094 (0.00181)	0.00109 (0.00216)	0.00115 (0.00244)	0.00116 (0.00244)
Wealth Effect Control	No	No	Yes	Yes	Yes
Individual Characteristics Control	No	No	No	Yes	Yes
Industry Effect	Yes	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes	Yes
Avg Prob (Buying New Stocks in an Industry)	0.0262	0.0262	0.0216	0.0196	0.0196
Observations	6,990,528	6,990,528	4,961,136	3,490,752	3,490,752

Table 4: Variation in Investor Sophistication

This table reports maximum likelihood regression results for probit regressions for subgroup sample. The sample is divided into four groups by self-reported experience(knowledge): Extensive, Good, Limited and None. The results are reported as marginal effects of independent variables. Each observation corresponds to a household/industry/year pair, regardless whether the household previously owned stocks in the industry or not. The dependent variable is based on a dummy variable coded one when a household purchases stocks not previously owned in the industry. “Experience dummy” is coded one if the household owns stocks in the industry in the past year. Other independent variables are dummy variables related to personal experience, industry average, portfolio size and individual characteristics variables. Both personal experience variables and industry average variables are based on the value of market-adjusted (experienced or industry average) return. “Good experience”(“Good industry”) is coded one when the market-adjusted experienced (industry average) return is greater than 0. “Top experience”(“Bottom experience”) is coded one when the market-adjusted experienced return is above (below) the 90th percentile (10th percentile) of the sample. “Top 1 industry”(“Bottom 1 industry”) is coded one when the industry average return is the highest(lowest) among the 10 industries. “Increase of Portfolio Size” is 1 denoting the size of household’s portfolio increases in the past year. Standard errors, shown in parentheses, are clustered by industry-year level. *10%, **5%, ***1% significance.

	(1)	(2)	(3)	(4)
Dependent Variable:	Buy New Stocks (not Owned in the Past Year) in One Industry			
Sophistication Group:	None	Limited	Good	Extensive
Experience Dummy	0.07743*** (0.01099)	0.07533*** (0.00491)	0.08738*** (0.00460)	0.11594*** (0.00830)
Good Experience (> 0)	0.01935* (0.01113)	0.01128** (0.00469)	0.01249*** (0.00378)	0.00563 (0.00721)
Top Experience (over the 90th percentile)	-0.01203 (0.01390)	0.00257 (0.00664)	0.00867 (0.00626)	0.00118 (0.00804)
Bottom Experience (below the 10th percentile)	-0.02573** (0.01307)	-0.00688 (0.00463)	-0.00763 (0.00548)	-0.00677 (0.00952)
Good Industry (> 0)	-0.01277 (0.01210)	0.00522 (0.00489)	0.00473 (0.00634)	-0.00526 (0.00859)
Top 1 Industry	0.02227 (0.02009)	0.03381*** (0.01094)	0.04296*** (0.01500)	0.06025** (0.02479)
Bottom 1 Industry	-0.00019 (0.01730)	0.01045 (0.01063)	0.02446 (0.01557)	0.02722 (0.02082)
Increase of Portfolio Size	0.01895*** (0.00566)	0.01352*** (0.00310)	0.02015*** (0.00281)	0.00978*** (0.00346)
Industry Effect	Yes	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes	Yes
Observations	43,580	143,570	97,440	10,080

Table 5: Variation in Portfolio Diversification

This table reports maximum likelihood regression results for probit regressions. The results are reported as marginal effects of independent variables. Each observation corresponds to a household/industry/year pair, regardless whether the household previously owned stocks in the industry or not. The dependent variable is based on a dummy variable coded one when a household purchases stocks not previously owned in the industry. “Experience dummy” is coded one if the household owns stocks in the industry in the past year. Other independent variables are dummy variables related to personal experience, industry average, portfolio size and individual characteristics variables. Both personal experience variables and industry average variables are based on the value of market-adjusted (experienced or industry average) return. “Good experience” is coded one when the market-adjusted experienced return is greater than 0. “Top experience” (“Bottom experience”) is coded one when the market-adjusted experienced return is above (below) the 90th percentile (10th percentile) of the sample. The independent variables also include the interaction terms between experience-related variables and portfolio diversification variables. The “Diversified” variable equals one when: either (1) the Herfindahl index¹⁷ is smaller than the median; or (2) the number of stocks in the portfolio is greater than 5, which is the median of the number of stocks owned by the households in the sample. The industry average return variables, wealth effect and individual characteristics controls are also included. The definition of these variables are noted in Table 2. Standard errors, shown in parentheses, are clustered by industry-year level. *10%, **5%, ***1% significance.

Dependent Variable: Diversified Portfolio Measured by:	(1) Buy New Stocks (not Owned in the Past Year) in One Industry Herfindahl < Median	(2) Num of Stocks > 5
Experience Dummy	0.06365*** (0.00972)	0.06163*** (0.00960)
Good Experience (> 0)	0.01392*** (0.00429)	0.01324*** (0.00360)
Top Experience (over the 90th percentile)	0.00240 (0.00527)	0.00256 (0.00387)
Bottom Experience (below the 10th percentile)	-0.01052*** (0.00371)	-0.00927*** (0.00281)
Experience × Diversified	0.00997*** (0.00226)	0.01362*** (0.00204)
Good Experience × Diversified	-0.00756** (0.00306)	-0.00853*** (0.00265)
Top Experience × Diversified	0.00229 (0.00546)	0.00254 (0.00412)
Bottom Experience × Diversified	0.00458 (0.00522)	0.00301 (0.00363)
Diversified	0.04277*** (0.00181)	0.05231*** (0.00198)
Industry Return Control Variable	Yes	Yes
Wealth Effect Control	Yes	Yes
Individual Characteristics Control	Yes	Yes
Industry Effect	Yes	Yes
Year Effect	Yes	Yes
Observations	777,630	777,630

¹⁷The Herfindahl index of the portfolio is defined as $\sum_k w_k^2$, where w_k denotes the portfolio weight allocated in stock k and $\sum_k w_k = 1$.

Table 6: The Impact of Personal Investment Experience on Future Purchases in the Experienced Industry vs Other Industries

This table reports maximum likelihood regression results for one probit regression. The results are reported as marginal effects of independent variables. Each observation corresponds to one household and one industry, regardless whether the household previously owned stocks in the industry or not. The dependent variable is based on a dummy variable coded one when a household purchases stocks not previously owned in the industry. Column 1 corresponds to purchase in the experienced industry, while Column 2 and 3 correspond to two other industries one of which is the most similar to the experienced industry while the other one of which is the most different. The most similar and the most different industries are selected by measuring the distance between the stock returns in that industry and those in the experienced industry. The distance is measured by averaging daily absolute difference between the stock returns of two industries. “Experience dummy” is coded one if the household owns stocks in the industry in the past year. Other independent variables are dummy variables related to personal experience, industry average, portfolio size and individual characteristics variables. The personal experience variables are based on the value of market-adjusted experienced returns. “Good experience” is coded one when the market-adjusted experienced return is greater than 0. “Top experience” (“Bottom experience”) is coded one when the market-adjusted experienced return is above (below) the 90th percentile (10th percentile) of the sample. The industry average return variables, wealth effect and individual characteristics controls are also included. The definition of these variables are noted in Table 2. Standard errors, shown in parentheses, are clustered by industry-year level. *10%, **5%, ***1% significance.

Dependent Variable:	(1)	(2)	(3)
	Buy New Stocks (Not Owned in the Past Year) in		
	the Same Industry	Most Similar Industry	Most Different Industry
Experience Dummy	0.06571*** (0.00937)	0.03103*** (0.00190)	0.01257*** (0.00105)
Good Experience (> 0)	0.01856*** (0.00312)	0.00249* (0.00129)	0.00188 (0.00137)
Increase of Portfolio Size	0.01482*** (0.00168)	0.01535*** (0.00214)	0.01226*** (0.00254)
Average Num of Stocks > 5	0.05685*** (0.00205)	0.06603*** (0.00280)	0.03713*** (0.00157)
log(Average Portfolio Size)	0.01539*** (0.00072)	0.01560*** (0.00108)	0.01166*** (0.00146)
log(Average Turnover Rate)	0.00851*** (0.00078)	0.00752*** (0.00157)	0.00721*** (0.00144)
Experience Variable Corresponding to the Industry in the Dependent Variable	-	Yes	Yes
Industry Average Return Corresponding to the Industry in the Dependent Variable	-	Yes	Yes
Industry Effect	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes
Observations	777,630	777,630	777,630

Table 7: The Impact of Personal Investment Experience: Variation in Categorization

This table reports maximum likelihood regression results for one probit regression. The results are reported as marginal effects of independent variables. The sample are the observations corresponding to the subindustries within the Manufacture industry (defined as Industry 3 in Fama-French 10 Industry). Each observation corresponds to a household/subindustry/year pair. The subindustry is classified by Fama-French 48 Industry. The dependent variable is based on a dummy variables coded one when the household purchases stocks not previously owned in one subindustry. Column 1 corresponds to the purchase in the experienced subindustry. Column 2 corresponds to a random subindustry (other than the experienced one) in the Manufacture industry. Column 3 corresponds to a random subindustry outside the Manufacture industry. “Experience dummy” is coded one if the household owns stocks in the subindustry in the past year. “Good experience” is coded one when the market-adjusted experienced return is greater than 0. “Increase of Portfolio Size” is 1 denoting the size of household’s portfolio increases in the past year. The three individual variables are created from the beginning-of-month position data. They denote average monthly number of stocks, average monthly size and average monthly turnover rate in the past year. Standard errors, shown in parentheses, are clustered by industry-year level. *10%, **5%, ***1% significance.

Dependent Variable:	(1)	(2)	(3)
	Buy New Stocks (Not Owned in the Past Year) in		
	the Same Subindustry	A Random Subindustry in Manufacture Industry	A Random Subindustry Outside Manufacture Industry
Experience Dummy	0.01256*** (0.00094)	0.00572*** (0.00083)	0.00486*** (0.00054)
Good Experience (> 0)	0.00148** (0.00067)	0.00158* (0.00094)	0.00010 (0.00075)
Increase of Portfolio Size	0.00263*** (0.00022)	0.00476*** (0.00032)	0.00349*** (0.00025)
Average Num of Stocks > 5	0.00945*** (0.00039)	0.01513*** (0.00039)	0.01111*** (0.00035)
log(Average Portfolio Size)	0.00246*** (0.00012)	0.00365*** (0.00013)	0.00306*** (0.00011)
log(Average Turnover Rate)	0.00181*** (0.00010)	0.00189*** (0.00014)	0.00181*** (0.00010)
Experience Variable Corresponding to the Subindustry in the Dependent Variable	-	Yes	Yes
Industry Average Return Corresponding to the Subindustry in the Dependent Variable	Yes	Yes	Yes
Industry Effect	Yes	Yes	Yes
Year Effect	Yes	Yes	Yes
Observations	1,309,032	1,309,032	1,309,032

Table 8: The Impact of Personal Investment Outcomes on Subsequent Trade Frequencies in an Industry

This table reports the results of Tobit regression of the number of trades in a industry during the subsequent year on the investor’s investment outcomes in that industry. Each observation corresponds to a household/industry/year pair. The observation is only included if the household previously owned stocks in that industry. The dependent variable is the number of trades the investors execute in one industry, including both buys and sells. The personal experience outcome variables are based on the value of market-adjusted experienced return. “Good experience” is coded one when the market-adjusted experienced return is greater than 0. “Top experience” (“Bottom experience”) is coded one when the market-adjusted experienced return is above(below) the 90th percentile (10th percentile) of the sample. The control variables include industry average variables, portfolio size, and individual characteristics variables. The definitions of these variables are noted in Table 2. Standard errors, shown in parentheses, are clustered by industry-year level. *10%, **5%, ***1% significance.

	(1)	(2)
Dependent Variable:	Number of Trades in the Same Industry in the Subsequent Year	
Good Experience (> 0)	0.62076*** (0.09937)	0.45660*** (0.08877)
Extremely Good Experience (over the 90th percentile)	1.86391*** (0.21580)	1.25045*** (0.13850)
Extremely Bad Experience (below the 10th percentile)	-0.05677 (0.09374)	0.10375 (0.13370)
Good Industry (> 0)	-0.08504 (0.19169)	-0.08349 (0.18384)
Best Industry	1.03377*** (0.37715)	1.44166*** (0.30588)
Worst Industry	0.64351 (0.42882)	0.59849 (0.38981)
Wealth Effect Control	No	Yes
Individual Characteristics Control	No	Yes
Industry Effect	Yes	Yes
Year Effect	Yes	Yes
Observations	385,552	225,142

Appendix: Details on Processing Data

I implement some restrictions to select the trading records related to the empirical analysis in this paper, and during the process of combining trading records data with SIC code or price data from CRSP or Compustat, I have to eliminate some observations with missing data. This Appendix details the steps I took to get the final subsample used in this paper.

Before combining with data from other datasets, I select the trading records through three steps. First, among the trades of various investments, I only retain those related to investors' direct investments on common stock. Second, I remove the observations of households if there are inconsistent buy/sell records and quantity records for their trades. For example, the trading activity is recorded as "B(uy)" (or "S(ell)"), while the quantity of that trade is recorded as a negative number (or a positive number). Third, I eliminate the observations of households if the households have trades including short sell, more specifically, if they once had negative cumulative shares on some stocks.

The next step is to combine the trading records with SIC code from CRSP and Compustat. As I described before, I get the Cusips for the stocks invested and match them first with CRSP, and then with Compustat if I can't find a match in CRSP. Only all the investments of the household could match with a SIC code either from CRSP or Compustat, the observations of this household will be selected into the final subsample used for the baseline regression.

Table 9: Number of Households Retained after Each Refinement

	Number of Households
With direct investments on common stocks	66,465
Without inconsistent buy/sell records and quantity records	62,554
Without short-selling trades	54,210
Without missing SIC code	47,793

In other analysis (variation in sophistication, time horizon or categorization), due to missing data of specific related variables, I have to further restrict the sample. Through robustness check, basic summary statistics and main results remain the same with different samples.