University of Michigan Retirement Research Center

Working Paper

WP 2006-125

Understanding Trading Behavior in 401(k) Plans

Takeshi Yamaguchi

MRRC

Project #: UM05-S2

"Understanding Trading Behavior in 401(k) Trading Plans"

Takeshi Yamaguchi
The Wharton School

September 2006

Michigan Retirement Research Center
University of Michigan
P.O. Box 1248
Ann Arbor, MI 48104
http://www.mrrc.isr.umich.edu/
(734) 615-0422

Acknowledgements

This work was supported by a grant from the Social Security Administration through the Michigan Retirement Research Center (Grant # 10-P-98358-5). The findings and conclusions expressed are solely those of the author and do not represent the views of the Social Security Administration, any agency of the Federal government, or the Michigan Retirement Research Center.

Understanding Trading Behavior in 401(k) Plans

Takeshi Yamaguchi

Abstract

We use a new database covering 1.2 million active participants to study trading activities in 1,530 defined contribution retirement plans. Descriptive statistics and regression analysis indicate some interesting trading patterns. First, we show that trading activity in 401(k) accounts is very limited: only 20% of participants ever reshuffled their portfolios in two years. Second, demographic characteristics are strongly associated with trading activities: traders are older, wealthier, more highly paid, male employees with longer plan tenure. Finally, we find that plan design factors, such as the number of funds offered, loan availability, and specific fund-families offered have significant impacts on 401(k) plan participants' trading behavior. Moreover, on-line access channels stimulate participants to trade more frequently, although they do not increase turnover ratio as much. We conclude that plan design features are crucial in sharing trading patterns in 401(k) plans.

Authors' Acknowledgements

I thank Olivia S. Mitchell for her support and guidance during the research process. I thank Stephen Utkus and Gary R. Mottola from the Vanguard Group for providing excellent data and suggestions. I am also grateful for comments from Prof. Terrance Odean, Tongxuan Yang, Sojong Park, Qingyi Song. Any error is my own.

1. Introduction

Individually-managed 401(k) accounts are now the most widespread investment form for retirement saving in the U.S. Currently, about one-third of all workers are enrolled in 401(k) plans, managing near $1.9 trillion in assets[2] worth about 20% of the U.S nominal GDP. Defined contribution schemes are also becoming more popular in other countries including Australia, Chile, Singapore, and Sweden where investment choices are offered to participants. Japan passed its Defined contribution pension Act in Oct 2001, and Korea also started a DC plan scheme in December 2005.

Although some theoretical financial research papers suggest that the individual trading behavior can influence the returns and volatility of the financial market (e.g. Bertola and Foresi 1995, Odean 1998), little is known about how people actually invest their 401(k) plan assets. Further, there is nothing known about how their trading patterns might influence retirement outcomes. Recent empirical research by Odean and his colleagues (Odean 1999; Barber and Odean 2000, 2001) provides evidence on trading activities and portfolio performance of individual investors, but these papers focus on the discount brokerage accounts, representing only a narrow and probably non representative subset of the universe of individual investors. In addition, because discount brokerage account holders are those who presumably like to trade, it is necessary to pay attention to self-selection bias to interpret their results. For defined contribution accounts, as far as we know, Agnew (2003) is the only empirical analysis focusing on the relationship between individual trading activities and their demographic characteristics in 401(k) plans. Consequently, nevertheless that it is necessary to confirm whatever previous results can be extended to other broader databases. We also need to investigate whether other factors beyond demographics also drive participants to trade their pension portfolios.

In this paper, we use an invaluable new dataset, provided by the Vanguard Group on more than 3 million 401(k) participants in over 2,000 plans, to illustrate how participant and plan characteristics shape trading behavior in individually-held voluntary retirement accounts. We examine the patterns of trading in 401(k) plans, focusing on who trades, how often, what their

[2] The latest estimate about 401(k) plans by Department of Labor is for plan year 1999, which was released in Summer 2004. In 1999, there are 335,121 401(k) type plans covering 38.6 million active participants with $1,790 billion in assets. EBRI/ICI (Employee Benefit Research Institute/Investment Company Institute) estimates 401(k) plan universe held $1,885 billion in asset at year-end 2003.

turnover patterns are, as well as how plan design and other plan features influence their trading activities.

The rest of this paper is organized as follows. In Section 2, we review previous literature on individual investor trading patterns. Section 3 describes the dataset. Section 4 presents descriptive statistics regarding trading behavior. In Section 5, we present our multivariate analysis methodology, hypotheses, and regression results. Conclusions appear in Section 6.

2. Previous Literature

Although there is a substantial theoretical literature regarding individual portfolio trading behavior, most of it has not focused on investment patterns in defined contribution retirement accounts. For instance, the conventional financial theory suggests that investors will trade when the marginal benefit of trading equals or exceeds the marginal cost (Grossman and Stiglitz, 1980). By contrast, overconfidence theory predicts that investors will trade to their detriment (Gervais and Odean, 2001). Recently, Mitchell, Muermann, and Volkman (2004) have found that anticipated regret regarding investment outcomes prevents a regret-prone investor from making extreme portfolio allocations, which would make him more likely to rebalance his account as compared to a more conventional rational investor.

To date, little empirical literature has explored trading patterns in retirement accounts, though there has been much research on interesting and unexpected patterns of trading in discount brokerage accounts. A set of studies by Odean and his co-authors (Odean 1999; Barber and Odean 2000, 2001) supports the overconfidence theory more than rational expectations framework with the existence of transaction cost. Using data on trading activities of 78,000 discount brokerage accounts, they computed the transaction cost, defined as the sum of bid-ask spread and commission fees. Comparing realized returns and the transaction costs, they provide evidence on trading behavior and investment performance of individual investors. For an instance, they conclude that people tend to trade too much; trading behavior appears hazardous to investors' portfolio performance (Barber and Odean, 2000); men appear to trade more than women because they are more overconfident (Barber and Odean, 2001).

In addition, King and Leape (1987) examine the changes in portfolio composition over the life cycle from the transaction cost approach. Using 1,978 Survey of Consumer Financial Decisions conducted by SRI (Stanford Research Institute), they studied the portfolio selection problem of over 6,010 U.S. households. They classify investment vehicles into "information-intensive assets" such as individual equity and other assets, and then they report that the ownership of "information-intensive" assets increases even after controlling changes in wealth, marginal tax rate and household characteristics. Therefore, the authors conclude that leaning and experience are important factors in lowering transaction costs overtime, leading to a positive relationship between investors' ages and "information-intensive" portfolio selection.

Empirical studies on trading in 401(k) plans are few in number, mainly because until now, analysts have not had good quality data on a wide variety of plan designs and participant outcomes. To our knowledge, Agnew et al (2003) focusing on a single large 401(k) plan is the only empirical work which explores participants' trading patterns and their portfolio selection problem. That analysis studied about 6,500 retirement accounts during April 1994-August 1998. At first, the authors examined the asset allocation and trading activities patterns using tabulations, and then they regressed equity allocation, turnover ratio and annual trade number on a set of demographic variables including age, sex, salary, marital status, and time in the plan. In addition, Agnew et al. also track participants over time, which makes it possible to investigate how equity allocation and trading behavior change as individuals' age and gain seniority on the job. It also allows their study to model the dependent variables as a function of common time effect. Finally, they conclude that most asset allocations are extreme (either 100% or 0%), and there is inertia in asset re-allocations; equity allocations are higher for males, married investors and for investors with higher earnings and more seniority on the job; equity allocations are lower for the elder people. They find only limited portfolio reshuffling, in sharp contrast to discount brokerage accounts.

Whether the results in Agnew et al. (2003) are generalizeable is in question. For example, it would be useful to confirm whether their finding is specific to that single plan, or rather whether it is a general phenomenon across plans. Moreover, when interpreting the common time-effect both on equity allocations and trading behavior, it is also necessary to take into account that before April 1994 (the beginning point of their database), participants were only able to invest in Guaranteed Investment Contracts (GICs). In addition, we also need to investigate whether other factors beyond demographics can also influence individual trading activities.[3] In particular, we seek to learn whether plan design is sharing participants' trading behavior.

As far as we know, our current investigation is the first comprehensive study on 401(k) trading activities across multiple plans. It is distinguished from previous studies in 2 ways: First, we study a database of over 3 million DC retirement accounts across above 2,000 plans, which will enable us to examine a larger subset of the 401(k) universe. Second, we can study trading patterns at the participant-level as a function of both participant and plan-level variables.

3. Database and Descriptive Statistics

In this section, we describe our new database. In Part A, we introduce the raw data and discuss its reprensentativeness. Part B summarizes investment options, industry, and plan-level information. Part C presents participant-level characteristics.

A. The Vanguard Database

This study uses a new database provided by the Vanguard Group (VG), contained information on more than 3 million DC plan accounts across 2,252 401(k) plans. Those plans included 984 funds in the investment menu[4] over the period Jan 2003 - Dec 2004. This is hierarchical database that consists of 5 levels: time, industry, plan, participant and trade.

At the industry level, the database codes every plan sponsor based on 4-digit NAIC[5] criteria, which have been aggregated into ten industries.[6]

At the plan level, the plan files record the major plan design features such as the number of funds offered, employer company stock as an investment option, and loan availability. In particular, this file enables us to figure out changes in investment menu for each plan, so we can evaluate the impact on participant level trading behavior.

---

[3] An analysis of just two 401(k) plans by Choi, Laibson and Metrick (2002) evaluates only the impact of internet access to trading behavior at *plan-level*, but does not look at trading patterns and the impact on *participant-level*. Benartzi and Richard Thaler (2001) consider a cross-section of plans, rather than a cross-section of individuals and investigates how allocations at the *plan-level* change as a function of the investment menu in the plans.
[4] Actually, company stocks and cash are also for part of p lans, for convenience, we just call every available investment choice as "fund".
[5] North American Industry Classification.
[6] 1) Agricultural, Mining, or Construction, 2) Transportation, Communications or Utilities, 3) Manufacturing, 4) Media, Entertainment, Leisure, 5) Wholesale & Retail Trade, 6) FIRE - Finance, Insurance, Real Estate, 7) Business, Professional and Non-Profit Services, 8) Education & Health, 9) Government, 10) Unknown.

The participant level files give us each participant's demographic characteristics and monthly balance amount as well as his portfolio allocations. The trade level file records trade directions (e.g. sell company stock to buy Money Market Fund) and trade amounts. In addition, we also obtained a separate file describing fund features such as fund classification (e.g. U.S equity, international equity or MMF).

Compared to previous empirical studies, this database is far more representative of the 401(k) covered population. For example, if we take the EBRI/ICI database[7] as a proxy of the 401(k) universe, we can compare some plan and participant-level features available both for EBRI/ICI and the VG database (insert Table1).

At the plan level, the total assets under management in the VG are over $146 billion at the end of 2003, around 20% of the EBRI/ICI amount. But the EBRI/ICI database represents only about 41% of the total 401(k) plan assets, so we conclude the VG database represents about 7% assets of the participant population. For the account number, 2.5 million participants hold 401(k) accounts in VG, one-sixth of EBRI/ICI data. Almost of the plans in EBRI/ICI data have a few participants, wherever the average plan scale in the VG dataset is about 4 times larger for both total balances and accounts per plan. It is interesting that at the participant level, the two datasets are very close in the terms of median in age, plan tenure and portfolio allocations.

For the trading analysis, we eliminate some observations with. At the plan level, we exclude three plan types of IRA, Deferred Compensation, and ESOP that do not fit into our 401(k) focus. At the participant level, we focus only on the accounts with continuous balance history (25 monthly balance records including Dec 2002) and continuous active plan status.[8] We exclude non-active accounts because the trading activities of retired participants maybe differ from those of active workers; further, we need 25 monthly balance records to compute turnover ratios. We also exclude some noise observations in the trade file such as the trades with only one-side record (only sale or only buy) or zero trade amounts.

[7] Employee Benefits Research Institute / Investment Company Institute. For further information about EBRI/ICI database, refer to "*Perspective*" Vol. 11 / No. 4A September 2005.
[8] For a discussion of plan status, refer to Appendix 1.

B. Investment Menu, Industry-level, and Plan Level Factors

After excluding inappropriate observations, our final sample size for the trading analysis is 1,186,554 accounts across 1,530 distinct 401(k) plans. This subsection reports the characteristics of the plans investment menu, industry and plan design features.

Over the sample period (Jan 2003 – Dec 2004), there are 691 unique funds ever selected as an investment option. The Vanguard Group has classifies those funds into seven primary types as: 1) Balanced Funds, 2) Bond Funds, 3) Equity Funds, 4) Money Market Funds, 5) Other Funds, 6) Unfunded Funds, and 7) VG Brokerage Option, of which 5) and 6) are stable-value funds. In addition, Equity Funds are categorized into five secondary types, which are 1) Aggressive Growth funds 2) Company Stock Funds, 3) Growth & Income Funds, 4) Growth Funds and 5) International Equity. Money Market Funds are also divided into two types as Investment Contract Funds and Money Market Funds. For convenience, we re-arrange all funds into three broader families defined as Equity, Fixed Income and Balanced Funds (the latter are hybrid investment vehicles with pre-determined target allocations in Equity and Fixed Income). Across the offered menu, almost 80% are equity funds and the other 20% are Fixed Income and Balanced Funds (insert Table 2).

Given the investment choice available in each plan, participants are allowed to alter their asset allocation patterns without any commission fees on a daily basis. It is worth nothing that all participants in a given plan will be required to reallocate their assets from one specific fund to the other available options, if their plan sponsor decides to drop a given fund from the menu. In this case, we call this type of trade as "Plan-wide trade".[9] On the other hand, participants are never required to allocate any asset to newly-added funds.

In Table 3 (insert Table 3), we summarize the industry level and plan-level characteristics of our sample. On average, plans offered about 17 funds as of Jan 2003, of which 10 were equity funds, 4 fixed income funds, and 3 balanced funds. There was a very wide range of plan design with 86 funds for maximum and one fund as minimum. It is also interesting to document that the average participant only holds a positive balance in 3.5 funds. In order to check whether menu change affects participants' trading activities, we also check whether plans have ever changed their investment option. Of the 1,530 plans, there were 513 plans which ever added new funds

[9] For further definition of plan-wide trade, refer to Appendix 2.

and 121 plans ever dropped at least one fund, and the rest 896 never changed their investment options.

Turning to the investment menus offered, most plans offered indexed equity funds, and international equity funds were available for over 90% plans. Although only in 15% plans (229) the employer company stock was a possible choice, because those plans are large, half of the participants could invest in their employer's share. In January 2003, about three quarters of the plans allowed participants to borrow from their DC retirement account, which covers 85% sampled participants.

We also review the distribution of life-cycle fund offering dummy. Life-cycle funds, also called as "life-stage" or "target-date" funds provide a mechanism to automatically reduce the proportion of the portfolio held in equities as an investor ages. In our sample, the Vanguard Group offers a set of "Target Retirement Funds", each of which has a target retirement year (2005, 2015, …, 2045). As the target year approaches, the investment management will gradually shift the funds from equities to bonds or MMF. In practice, the allocation changes are not directly linked to any particular investor's age, but those funds are marketed to those who plan to retire in time close to the target date of the fund. The idea is that investors who do not want to or are unable to spend money and time to take care of their asset allocation, will more easily manage their retirement assets. In our sample, there are 744 plans (49%) providing this kind of investment vehicle.

C. Participant Characteristics

Next, we present descriptive statistics of participant level data. Demographic characteristics of participants are shown in Panel A of Table 4 (insert Table 4). The majority of our sample is male (47.5%), 26.7% of the accounts are held by females, and the reminder (25.8%) is coded as missing. Compared with Agnew's sample, our sample accounts show an older mean age (43.5 vs. 40.0 years old) with a wider range (our maximum is 83 years old and the minimum is 16). Household income class (coded as 1, 2,…,11) is imputed by Claritas[10] for 2003 using participants ZIP code. We converted these income classes into income levels using a lognormal distribution. In 2003, the average household income is $88,000, somewhat higher

[10] Claritas is a marketing information resources company which provides inome by ZIP code, for detail refer to http://www.claritas.com.

than Agnew et al. (2003)'s sample mean ($69,000). For financial wealth, we use IXI Wealth class variables provided by IXI Corporation.[11] and we convert it into mean dollars of each class. The sample mean value in 2003 was $75,000.

Panel B of Table 2 presents some interesting participant-level variables. Employee Contribution Ratio, defined as the Employee-contributed Balance divided by the Total Balance, is important not only because it reflects the plan-match ratio, but also it expresses participants' attitude towards contribution to their retirement accounts.[12] In addition, it is also possible that attitudes toward "own money" and "house money" are different, which could affect participants trading behavior. In January 2003, the mean of Employee Contribution Ratio is 57.5%, meaning that only 57.5% of the total balance of the average 401(k) account is contributed by the employee himself, his employer contributes the other 42.5%.

On-line trading has been proposed as a cause of excessive trading, Metrick et al. (2002) and several previous studies show that internet trading channels increase individual investors' trading frequency in discount brokerage accounts and 401(k) plans.[13] We generate a web-access dummy variable, defined as equal to1 if a participant holds a web access to his account, and 0 otherwise. Some 37.4% of our sample accounts are able to do web trading.

4. Evidence of Trading Activities

In this section, we provide the definitions regarding trading activities, present descriptive statistics, and then summarize some trading patterns.

A. Measuring Trading Activities

All trade orders issued by a 401(k) plan participant within the same day will be executed at the end price (3:00pm) for each fund. Accordingly, we define a ***Trade*** as an asset allocation change in a day, which consists of both a sell record and a buy record of at least one fund and a positive trade amount become recorded. For each trade, we compute the total sale/buy allocation

---

[11] IXI is a private company which provides wealth measures by ZIP code, for detail, refer to http://www.ixicorp.com.
[12] This ratio might be different across accounts in a given plan. e.g, given 15% as the upper limit for total compensation, 6% as upper limit as employer contribution with 100% match ratio, a participant with 6% employee contribution has different Employee Contribution Ratio with a person with 9% employee contribution.
[13] As the example of empirical studies based on discount brokerage accounts, refer to Barber and Odean (2001), as a example of 401(k) plan studies, refer to Choi, Laibson and Metrick (2002).

change amount across all funds regarding this trade, and then we define the average of the sale and buy amount as ***<u>Trade Amount</u>***. This calculation can be written as below:

$$trdmt\_out = \sum_{i=1}^{n} trdmt\_out\_fnd_i$$

$$trdmt\_in = \sum_{j=1}^{m} trdmt\_in\_fnd_j$$

$$trdmt = \frac{trdmt\_out + trdmt\_in}{2}$$

trdmt_out: total sale amount across all sold funds;

$trdmt\_out\_fnd_i$: sale amount of fund i;

n: number of sold funds;

trdmt_in: total buy amount across all bought funds;

$trdmt\_in\_fnd_j$: buy amount of fund j;

m: number of bought funds;

trdmt: trade amount

If an individual investor ever exchanged his asset allocation at least once from January 2003 to December 2004, then we define this participant as ***<u>Trader</u>***, otherwise, we considered him as a ***<u>Non-trader</u>***. Each participant's total trade times is defined as his ***<u>Trade Number</u>***. Because plan-wide trades are not due to investors' choice but rather driven by the employer, we do not count those allocation changes as trades; therefore, if a participant only has plan-wide trades covering the sample period, we include him as a non-trader. To measure trading magnitude, we use the ***<u>Turnover Ratio.</u>*** This is defined as the ratio of total trades amount over participants sample average balance, defined as the average balance of sample beginning (December 2002) and sample end (end of Dec 2004). This calculation could also be re-written as below:

$$trdmt\_smpl = \sum_{l=1}^{k} trdmt_l$$

$$fndtotasset\_avrg = \frac{fndtotasset_t + fndtotasset_{(t-1)}}{2}$$

$$Ratio\_turnover\_smpl = \frac{trdmt\_smpl}{fndtotasset\_avrg}$$

trdmt_smpl: total trade amount across total trades in sample period;

$trdmt_l$: trade amount of trade l;

k: total trade number in sample period;

$fndtotassets_t$: total assets in the end of period t;

fndtotassets_avrg: average total assets;

Ratio_turnover_smpl: turnover ratio.

B. Plan-level Trading Activities

Previous research suggests that offering a large number of funds provides diversification opportunities for knowledgeable participants. Thus can lead to increase trading probabilities; on the other hand, too much variety in the investment menu may cause information overload problem (Agnew and Szykman, 2004): a lot of choices also potentially limit 401(k) participants' trading activities.

It is also possible to examine how investment package affects participants' trading behavior from an efficiency perspective. Given an efficient investment frontier, for example, maybe an indexed equity fund, a rational investor maybe only need to put his money in that fund, and he may not need to select individual assets to create his own efficient portfolio.

As described in section 3, industrial recommendation of life-cycle funds suggests that this type of investment vehicle makes asset management easier. Designed for one-stop investment shopping, life-cycle funds distribute investors' money among various other funds, changing their asset allocation as they enter different life stages. Of course, rational and engaged investors may find such kind of investment vehicles redundant. However, for investors who may not know their risk tolerance, the overwhelming number of options may seem like more of a curse than a blessing, therefore, life-cycle funds can make portfolio selection problem easier: automatic adjustment mechanism makes them maintenance-free for investors.

Another industrial story suggests that offering international funds provides international arbitrage opportunities, and thus, increase investors' trading probabilities. This may be particularly appealing to speculative transactions. For example, if one expected the value of Japanese Yen to increase enough to cover transactions cost tomorrow in the eastern U.S. time zone, just buying tonight in Tokyo money market, and then selling tomorrow in N.Y. market, may the investor make money.

Table 5 presents our data on trading activities sorted by plan-level variables. Clearly, plan design factors, as well as with participants' financial literacy, are very important on 401(k) participants' asset allocation behavior (insert Table 5).

Our data shows evidence to support those theoretical implications. For example, the plan-level trading propensity, which is defined as the ratio of trader to active participants per plan, has a hump shape associated with the number of funds offered. It increases from 14.4% in plans offering 1-5 funds to 17.7% and 21.5% in groups with 6-10, and 11-17 investment choice respectively, and then hits its peak of 23.6% in plans with 18-30 funds. Finally, it starts to decrease to 19.6% as a plan offers more than 30 funds. This pattern may suggest that adding new funds to investment menu does increase participants' trading probability due to more diversification opportunities, but this effect decreases marginally and there exists an upper bound.

Trading propensity in plans offering international equity funds or employers' company stock also is significantly higher than those plans without those types of investment choice. More than 20% of participants in an international equity available plan have ever reallocated their portfolios in our sample, which is 3.5% points higher than a plan lacking international equity funds. Company stock availability also contributes to activate exchange activities: there is 3.9% point spread of trading propensity between plans offering company stock and those not.

There are no obvious different trading patterns between plans offering life-cycle funds, indexed equity funds, and loan availability with those plans that do not have those options. In order to investigate if these variables influence participants' trading activities significantly, we need to analyze our data using a multivariable approach.

C. Participant-level Trading Activities

Table 6 presents the participant level trading activities, which are measured as 1) trading probability, 2) trade number, and 3) turnover ratio. Only 20.5% of the 401(k) accounts ever altered their portfolio allocations over two-year, which indicates trading activities in individual-managed retirement accounts are very limited. This result is also consistent which the statistics in Agnew et al. (2003) single plan study, which showed about 88% participants have no trades. In our data, conditional on trading, the average sample turnover ratio reached almost 90%, which means on average a trader changed his asset allocations by 90% over two-year period (insert Table 6).

Transaction cost theory tells us that investors will change their asset mix if when the marginal benefits from trading is no less than the marginal cost. Given fixed transaction costs, it will be optimal for investors to relocate their assets infrequently with a discrete amount. Transaction costs could be classified into explicit and implicit costs. The explicit costs are always expressed in a monetary form such as bid-ask spread, commission fees in case of common stocks or trustee fees in the case of mutual funds. On the other hand, implicit costs may take the forms of opportunity costs such as time spent on information collection and information analysis. Because trading is free for all of the sampled accounts, thus, in our case, implicit costs are supposed to be the only transaction cost.[14]

The implications of the transaction cost theory can also be captured by investors' characteristics. For instance, as King and Leape (1987) have argued, learning, experience and knowledge can lower the fixed transaction costs and thus lead to increase trading probability. This suggests that elder participants in knowledge-intensive industries (e,g, Finance, Insurance and Real Estate) are likely to reshuffle their portfolio more frequently with bigger magnitude. Meanwhile, longer plan tenure not only capture investors' degree of knowledge on their plan and trading experience, and also indicate bigger balances, which will increase the likelihood of trading.

Table 6 presents the trading probability for each group sorted by participants' demographic features. We see that men are more likely to trade their assets significantly. Nearly one quarter (24.1%) participants have ever changed their portfolio allocations over the two- year period, but only 17.9% among females. These unbalanced trading probabilities are consistent with the overconfidence hypothesis and the empirical results in Barber and Odean (2001): men are more overconfident than women, so men are more likely to trade. Of course, the sharp contrast between males and females could also be explained by the transaction cost framework[15]. That is females often have lower salary, smaller financial wealth and less investment literacy compared to men.

The trading probability increases as participants get older for almost all age ranges. For investors under 35, only 16.9% participants have trading experience, in 36-45 years old group, this ratio goes up to 18.9%, and then it climbs 22.8% in 46-54 group, and it finally hits its peak

[14] Of course, there is no free lunch. Ultimately, the transaction costs must be shared by mutual fund companies, plan sponsors, and participants in other forms.
[15] Refer to experiment study by Agnew (2004).

of 25.1% in 55-64 group. This increasing trend fits the transaction cost theory described above, and the sharp decrease in trading among the over 65 years old group can also be rationalized: the portfolio allocation should have been exchanged to relative safe assets, thus reduce the motivation to rebalance their assets to safer options.

Trading probability has a monotonically increasing trend with participants' plan tenure. The trading probability jumps from 17.4% in the participant group with less than 5 plan years, to 31.7% for investors who have already stayed in the same plan over 25 years. Obviously, this pattern is consistent with the theoretical implication as described above: the longer an investor stays in the same plan, the better he understands the plan, and the more experience/knowledge he has, which will increase his trading likelihood. On the other hand, presumably, longer plan tenure also usually implies a bigger plan balance, which indicates possible substantial benefits from trading activities, thus motivates the investor to exchange his asset allocations.

Wealthier investors and participants with higher salaries are also more likely to trade. We categorized our participant level observations as low income/wealth group if household income/IXI wealth was under the sample median, and those who were above median as the high group. The probability of investors in the high income group is 23.1% significantly, which is nearly 6% points higher than their low income counterparts. People in the high wealth group are likely to relocate their assets with a probability near one-quarter, but the poor participants did less than one-fifth of the higher.

Our finding is consistent with those results reported in previous studies: traders are older, wealthier, more highly paid, male employees, who have higher balance in their accounts. People in the financial industry also trade more frequently than in other industries.

Trade numbers are drawn from a very skewed distribution. As shown in Table 7, 79.5% accounts have zero trade in two years. Even among the traders, over half had only 1 trade. Roughly speaking, our 401(k) sampled participants altered their asset allocations 0.3 times in one year; that is to say they only trade once every 3.33 years on average. However, there is another extreme group who are trading more than 50 times over 2 years (insert Table 7).

Trade number patterns associated with participants' demographic characteristics are similar to trading probability. Males (0.75 trade/2 years) traded twice as frequently as female (0.36 trade/2 years), trade numbers increased with age until 65 years old, stayed in the same plan longer, earned more salary, and became wealthier. Another interesting pattern is that the trade numbers for participants with on-line accounts was 5.5 times (1.23 trades/ 2 years) higher than those lacking web access (0.22 trades/2 years). Of course, participants might open their on-line accounts because they want to trade. On the other hand, this sharp contrast still indicates that internet access availability does permit actual trading.

Turnover ratios correlate with trade numbers. On average, participants changed the composition of only 18% of their assets, but conditional on trading, the average turnover ratio for sample traders climbed to 89.7%. When we sort turnover ratios by participants' demographic characteristics, we confirm portfolio exchangers were likely to be men, older, high-income, wealthier, and had longer plan tenure investors compared their female, younger, low-income, lower wealth and shorter plan tenure counterparts. However, on-line access did not increase the sample turnover ratio as dramatically as it did on trade numbers, maybe because internet channels only convince people to trade more frequently but reduce the trade amount per exchange compared with the case if they had traded with a conventional way. This is also consistent with the empirical study results in Choi, Laibson and Metrick (2002), which is based on two large 401(k) plans plan level trading records.

All of these statistics are consistent with the single plan examined by Agnew et al. (2003). We conclude that trading activity is extremely limited in 401(k) accounts.

5. Regression Analysis

This section summarized hypotheses to test, describes our regression methodology, presents regression results and then compares them with the conclusions from other empirical studies.

A. Hypotheses

Table 8 presents independent variables associated with the hypotheses we seek to test, and shows the predicted signs in the regression analysis (insert Table 8). We can summaries it as follows:

1) Transaction Cost Hypotheses: Trading will be more prevalent among the people in financial literacy intensive industry like Finance, Insurance and Real Estate for industry level; older, highly paid, longer-tenured, wealthier, web-registered participants are more

likely to reshuffle their assets. Offering life-cycle funds can make portfolio selection easier, thus, it would reduce trading propensity.

2) Overconfidence hypothesis: Because men are more confident than women, trading will be more prevalent among males more than females.
3) Diversification Hypothesis: the more funds offered at the plan level, the more likely people will trade; as international funds provide arbitrage opportunities, international funds may boost trading likelihood; offering company stock diversify the investment menu, thus increases the trading probability.
4) Efficient investment frontier hypothesis: Assuming indexed funds span the efficient investment frontier, offering indexed funds will decrease investors' trading patterns.
5) Information Overload and Financial literacy Hypothesis: the information overload hypothesis proposes that more choice in the investment menu will reduce people's trading; females are often low salary, less financial knowledgeable, thus, females trade less.

B. Regression Methodology

For the regression analysis, we focus on three dependent variables: sample trader, trade number and turnover ratio. They measure the trading probabilities, trading frequency and trading magnitude, respectively.

We categorize the independent variables into three levels including five sub-classes: industry, plan, and participant levels. We do not control on time because two year is not long enough to estimate time effects. The specific independent variables are shown as below:

1) *Industry Level:* Industry Dummy. We code nine dummy variables based on the VG classification with the reference industry being manufacturing.
2) *Plan Level:* We divide plan-level explanatory variables into a) plan design features, and 2) other plan level variables
    a. Plan design variables: Number of funds offered, its square, and the change during sample period. Life-cycle funds dummy, International equity funds dummy, Company stock dummy, Indexed equity or/and fixed income funds dummy and loan availability dummy.

b. Other plan level variables: we adopt the natural log of the total plan assets as a proxy to measure scale economies.

3) *Participant level:* Participant-level independent variables are split into two sub-classes.

a. Demographic characteristics: sex, age, plan tenure, household income and wealth as well as missing dummies for each variable.

b. Other account-level features: On-line access dummy, employee-contribution ratio.

The regressions use a Generalized Linear Mixed Model (GLMM) framework. The reason for Generalized Linear Model (GLM) part results from the features of our dependent variables: Sample traders is a binary variable (0,1), trading number is a non-negative integer, and turnover ratio is a non-negative real number. The GLM relaxes the normality assumption required for a conventional regression model, and only requires the underlying distributions for dependent variables belong to the exponential family.[16] In addition, by mixing or compounding two distributions, GLM also enables us to handle models with more complicate underlying distributions.[17]

We also adopt a mixed error structure due to the hierarchical structure of our independent variables, which would generate serial correlation and heteroscadasticity problem. The multi-level regression approach provides extremely flexible variation for the error term structure, which enables us to test random effects, fixed effects and robust standard errors models.

In our analysis, we use the following regression model:

Linear Regression: $y_{i,k,m}{}^{*} = X_{i,k,m}'b + Z_{k,m}'g + W'_{m} d + w_{m} + u_{k,m} + e_{i,k,m}$

Linear Predictor: $m = E(y_{i,k,m}{}^{*} \mid X, Z, W)$

Link Function: $h = h(m)$

where $y_{i,k,m}{}^{*}$ is the latent variable for individual i in plan k whose sponsor belongs to industry m, X, Z, W indicate participant level, plan level, and industry level explanatory variable vectors; β, γ, δ are parameter vectors associated with X, Z, W. The error term consists of three parts: 1) a industry-variant term w_m, 2) a plan-variant term $\upsilon_{k,m}$, and 3) a disturbance term $\varepsilon_{i,k,m}$. Here, we control on industry-variant term w_m by using industry dummies; we also assume that the plan is the independent analysis unit, allow serial correlation and heteroscadastity within the same plan,

---

[16] Some usual examples are Gaussian, Inverse-Gaussian, Poisson, Gamma, and Binomial.

[17] It is well known that Negative Binomial is the mixture distribution of Poisson and Gamma distribution.

and then compute the cluster-robust standard error.[18] Thus, the new error term can be re-arranged as $x_{i,k} = u_k + e_{i,k}$, with $e_{i,k} \sim N(0, s_1^{\ 2})$ and $u_k \sim N(0, s_2^{\ 2})$, but the two terms are not necessary to be independent. Therefore the re-arranged error structure may be written as[19]

Error structure: $Cov(x_{i,k}, x_{j,l}) = 0$ for $k \neq l$

$$Cov(x_{i,k}, x_{j,l}) \neq 0 \text{ for } k = l$$

C. Regression Results

1) Sample Trader

With the GLMM approach, we can rewrite our regression function for Sample Trader according to a usual Probit regression model:

$$y_{i,k,m}^{\ *} = X_i'b + Z_k'g + W'_m d + x_{i,k,m}$$

$$m = E(y_{i,k,m}^{\ *} \mid X, Z, W)$$

$h = \Phi^{-1}(m)$, where Φ^{-1} is the inverse CDF of Normal Distribution

Table 9 presents the regression results. The estimator indicates a very high joint significance of the independent variables. The pseudo-R^2 is 10.9% and the concordant percent is 72.7%, which implies the specification validity of the whole model. Compared a model that include only demographic factors as in Agnew et al. (2003), our current model also shows a significant advantage. This evident in the absolute difference of the likelihood ratio between the two models, but it also appears in the improvement of pseudo-R^2 and concordant percent, which increase 8.4% points (10.9%-2.5%=8.4%) and 12.0% (72.7%-60.7%=12.0%) respectively (insert Table 9).

Our most remarkable finding is that most of the plan level variables affect investors' trading likelihood significantly. Trading probability rises as the number of funds offered grows as well as the number of fund changes. However, the sign of the square of number of funds offered is negative. From the marginal effects regarding the number of funds and number of

[18] Monte Carlo simulation results suggest the biggest cluster should not larger than 5% of total observations. In our case, 59,350 participants are enrolled in the largest plan, which is just 5% (59350/1186554=5%).
[19] Theoretically, it is also possible to build multi-level regression with multiple random effects: industry and plan. However, those computation tasks are beyond a usual computer capacity even using the most time-efficient optimization technique.

funds squared (+0.000561 vs. -0.00005), we can easily calculate the positive effects reach its peak at the 54th fund. This fact suggests that our results support the diversification hypotheses: adding new funds to the investment menu does stimulate participants to alter their portfolio allocations, although this effect diminishes as the total number of funds offered grows.

Offering a loan makes investors less much to change their asset mix with marginal effect of 2.2%. On the other hand, company stock availability encourages participants to change their portfolios. Meanwhile, investors do make their trading decisions depending on whether their plan sponsors offer an indexed equity fund. Controlling on other factors, a 401(k) plan participant is 44% less likely to trade if his plan offers an indexed equity fund, compared to a plan lacking such a fund.

We also find that offering Life-cycle funds reduces investors' trading probability because of its automatic risk adjustment mechanism, and an international equity fund encourages participants to change their asset allocations by 3.7% due to the potential international arbitrage opportunities.

Participant level independent variables confirm to those regarded in previous studies.

Demographic characteristics such as sex, age, and plan tenure positively correlate with investors' trading probability with consistent learning and experience hypotheses. Ten more years of age boosts the trading probability by 2.1%; ten more years of plan tenure arises it by 3%. Men alter their 401(k) portfolio composition 4.65% more than females.

Household income and financial wealth also significantly affect investors' trading probability. A 1% increase in household income raises people's trading probability by 2.9%, and the increase in the ratio of financial assets over household income also correlates positively with trading probability.

2) Trade Number

In the regression function for trade number, we assume the dependent variable is drawn from Poisson distribution, and the parameter of Poisson distribution is from a Gamma distribution. As a result, the trade number is assumed to be distributed as Negative Binomial, which is the mixture of Poisson and Gamma. In GLMM estimation, we use a log link function. In fact, we can interpret those statistical procedures as a method to capture individual heterogeneity in an econometric model.

Suppose the trade number is distributed as Poisson with the probability density function $\Pr(Y_i = y_i \mid X_i, Z_k, W_m) = \frac{e^{-l_i} l_i^{y_i}}{y_i!}$, in which λ is the Poisson distribution parameter. For the mathematical convenience, we also assume λ is distributed as Gamma with the density function $\Pr(u_i) = \frac{q^q}{\Gamma(q)} e^{-qu_i} u_i^{(q-1)}$, where Γ indicates the Gamma function with parameter θ and u.

If we add another individual heterogeneity term ζ into the error, a usual Poisson Regression with cross-individual heterogeneity could be written as:

$$\log(m_i) = X_i' b + Z_k' g + z_i = \log(l_i) + \ln(u_i) = \log(l_i u_i)$$

Mixing Poisson and Gamma distribution, we can obtain the density function unconditional on heterogeneity as:

$$\Pr(Y_i = y_i \mid X_i, Z_k) = \int_0^\infty \frac{e^{-l_i u_i}(l_i u_i)^{y_i}}{y_i!} \frac{q^q}{\Gamma(q)} e^{-qu_i} u_i^{(q-1)} du_i$$

Integrating respect to u_i, it is easy to get the closed form as the following, which is the familiar format for Negative Binomial distribution.

$$\Pr(Y_i = y_i \mid X_i, Z_k) = \frac{\Gamma(q + y_i)}{\Gamma(y_i + 1)\Gamma(q)} \left(\frac{1}{1+a}\right)^q \left(\frac{a}{1+a}\right)^{y_i}, \text{ Where } a = \frac{l_i}{q}$$

$$E(y_i \mid X_i, Z_k) = qa = l_i$$

$$Var(y_i \mid X_i, Z_k) = qa(1+a) = l_i\left(1 + \frac{1}{q} l_i\right)$$

In our regression function, we need to estimate parameter vectors, β,γ,δ in the Poisson part, and θ, the parameter in the Gamma part, which is also called overdispersion parameter.

Table 10 presents the estimator results. As before with the trade propensity function, our independent variables work significantly well to explain the dependent variable (insert Table 10).

For the plan-level variables, offering company stock and international equity funds are the main plan design features that stimulate participants to trade more frequently; these increase trade number by 23% and 17% respectively. Offering more funds significantly encourages investors to exchange more frequently, but the magnitude of the effect is relatively small (one more fund offered causes 1.1% rise in trade number).

Demographic characteristics also return their explanatory power to interpret individual investors' trading number as appeared in trade propensity function. Again, male, higher paid, richer, and longer plan tenured participants trade more frequently than their counterparts. Another interesting result is that individual on-line access motivates investors to exchange their asset mixture significantly more and with a very high magnitude.

In addition, the overdispersion parameter is estimated to be 4.78, which is significantly not equal to zero and suggests individual heterogeneity can not be captured by demographic characteristics alone.

3) Sample Turnover Ratio

Table 11 presents a Heckman second-stage OLS regression results for the turnover ratio. We adopt the Heckman second-stage regression rather than a conventional censored regression method because we want to explicitly decompose the explanatory power of independent variables into two parts: the indirect effects through trading probability, and the direct effects conditional on trading (insert Table 11).

The coefficient regarding the inverse Mills ratio from the 1st-stage Probit regression, is estimated to be 0.75, which suggests all of the independent variables jointly explain the turnover ratio indirectly through increasing trading probability.

At the plan-level, offering international equity funds and offering indexed equity funds are still correlated the turnover ratio as seen in the trading propensity and the trade number functions. However, the number of funds offered and its square changed their signs suggesting that traders tend to exchange their 401(k) assets a subset of funds, and do not intend to extend their portfolio variety.

D. Implication

So far, we have followed the analysis framework of Agnew et al. (2003) by investigating the trading patterns in 401(k) retirement accounts. We confirm the fact that the trading activity in 401(k) retirement accounts is extremely limited: only 20% participants ever altered their portfolios within a two year period. Demographics hence can explain trading patterns in the ways that supports the theoretical implications.

However, our analysis is distinguished with previous studies in several points.

First, we find that plan design can explain trading behavior. Offering indexed funds and loans significantly reduces participants trading, while offering international equity funds and employer' company stock gives investors more motivation to trade assets. The more funds the plan sponsor offers, the more actively participants trade, but that effect increases at a decreasing rate.

Second, instead of conventional Tobit regression, we use Heckman 2-stage regression method for sample turnover ratio analysis. Therefore, we explicitly show the direct and indirect effect from the same independent variable. For example, even controlling on traders (i.e. controlling on the indirect effects through trading probability), plan design still retain its significant power to explain the traders' turnover ratio.

It is easy to obtain the intuitive implications for plan sponsors and participants from our finding. Table 12 summarizes the marginal effects on 401(k) participants' trading behavior (insert Table 12).

As described above, although there is no any explicit transaction fee for trading 401(k) portfolios, ultimately, the transaction costs must be shared by mutual fund companies, plan sponsors, and participants in other forms. Therefore, the inertia trading activity in 401(k) accounts directly indicates cost saving for the plan sponsor if it shares part of the transaction costs. However, it is not clear whether the limit trading behavior is positive or negative for participants before examine the trading performance: maybe limit trading make investors pass over some better investment opportunities, it is also possible that limit trading protects investors from rather worse investment outcome.

On the other hands, it may be necessary to take into account the impact on participants' trading activity from the plan design factors when an employer designs its' 401(k) plan. For example, the most efficient way to minimize the trading probability, the number of trading, and the turnover ratio is to offer an indexed equity fund. In fact, more than 95% of our sampled plans hold at least one such fund in the investment menu. Meanwhile, offering international equity funds, employer company stock, and brokerage options is the effective way to diversify the investment options, and thus attract the participants. In order to control the trading probability to a pre-determined target, offering life-cycle funds can offset the positive diversification effects by adding 13 new funds to the investment menu.

Furthermore, significant explanatory power of participants' characteristics also requires the plan sponsors to pay attention to the demographic structure of the participants. For example, in an extreme case, compared to a 100% female company, a 100% male company needs to offer loan availability and life-cycle funds to balance out the overconfident effects of men.

Overall, our analysis provides the evidence that both participants and plan sponsors need to be aware of the behavior of opposite part. An employee seeking to maximize his net 401(k) saving needs to pay attention to the plan design features, because they significantly affect participants' trading behavior, and then their investment performance; an employer as fiduciary also need to take into account the impact on participants' trading behavior when it designs its 401(k) retirement plan.

6. Conclusion

Our analysis of 1.2 million 401(k) plan participants shows relatively little overall trading in DC retirement plans, with substantial heterogeneity.

We find that only 20% participants ever reshuffled their portfolios over a two-year period, indeed, most of the people never trade. Even when people do trade, it is rare and over half traders only trade once. Like previous studies, we find demographic characteristics can significantly explain trading activity: traders are older, wealthier, more highly paid, male employees with more money in their accounts. We also find plan design factors do have significant impact on 401(k) plan participants' trading behavior: participants trade more in plans with more investment choice, company stock, and international equity funds, and offering indexed funds, life-cycle funds, and loan availability reduce people's portfolio turnover. Finally, on-line access stimulates participants to trade more frequently, although they do not increase turnover ratio as much.

Our analysis also shows that trading activities in 401(k) plans differs from those in discount brokerage accounts. Compared to the results of Odean and his co-authors, individual investors seem likely to trade their asset through discount brokerage account rather than retirement saving: investors trade 1.44 times annually in the discount brokerage accounts but only trade 0.30 time in the 401(k) accounts. This sharp contrast maybe comes from 1) self-selection problem in discount brokerage accounts, and 2) limited choice in investment menu.

Maybe one possible direction for future research is to check the relationship between 401(k) trading and portfolio returns. In our analysis, we only use the characteristics of participants, plans, and industries, as independent variables, and we do not take into account the market environment. By computing how trading influence investors' portfolio performance, maybe we would deepen our understanding of the trading behavior in 401(k) plans.

Another reason for the inertia trading activity of 401(k) accounts perhaps comes from the efficiency problem of investment options. Financial economics theory teaches us that a rational investor will always rebalance their assets in order to allocate their portfolio in the efficient frontier associated with their utility functions with the competitive market assumption. However, in the 401(k) context, the investment menu offered by plan sponsors may not span the efficient frontier[20]; consequently, the potential inefficiency may have an impact on trading activity.

It deserves to note that our analysis only covers a short sample period of 2003-3004, which was an extremely bullish period for the U.S stock market. We need to extend the trading sample including the future period. Ideally, we want to include bearish market environment and check the robustness of the result in this paper.

Finally, in the future research, we plan to make a more detailed time-series approach to extend our future analysis. In particular, given the crucial role of the plan design features in participants' trading behavior, we need to evaluate how the plan design change influences employees' portfolio selection and investment performance in their retirement saving account, and therefore provides the policy implications to fund managers, record keepers, plan sponsors and policy makers.

[20] For example, Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake (2002) find about 62% 401(k) plans do not offer adequate investment choice.

Table 1: Descriptive Statistics and Comparison Data

	VG[1]	EBRI/ICI[2]	Comparison	
	(A)	(B)	(A/B)	(A-B)
Panel A: Selected Plan Features				
Total Plans	1,942	45,152	4.3%	
Total Accounts	2,525,268	15,047,358	16.8%[4]	
Total Assets ($ million)	146,944	775,984	18.9%	
Accounts per Plan	1,300	333	390%	
Assets per Plan ($ million)	76	17	440%	
Panel B: Selected Participant Features				
Age (median, year)	44.2	44		0.2
Plan Tenure (median, year)	8.7	7.0		1.7
Assets per Account ($)	58,189	51,569	113%	
Balanced Funds (%)	12.9%	9.5%		3.4%
Fixed Income Funds[3] (%)	26.1%	27.4%		-1.3%
Equity Funds (%)	60.8%	61.0%		-0.2%

1.VG: Vanguard Group

2.EBRI/ICI: Employee Benefit Research Institute/Investment Company Institute

3.Fixed Income Funds=Bond Funds+Money Market Funds+Guaranteed Investment Contracts (GICs)+Other Stable Value Funds

4.EBRI database represents 41% of the total 401(k) plan assets, which is about $1.9 trillion. Therefore, VG database represents about 6.9% of the universe.

All of the numbers are as December 2003.

Source: the Vanguard Group, "Perspective" Vol. 11 No. 4A, EBRI/ICI September 2005

Table 2: Investment Menu and Fund Families

Fund Family	Number of funds	(%)	Allocation as Jan 03
Balanced Funds	*43*	*6.22*	*12.95*
Balanced funds	42	6.08	12.81
Brokerage option	1	0.14	0.13
Fixed Income Funds	*100*	*14.47*	*31.05*
Bond funds	40	5.79	9.27
Money Market Funds	57	8.25	21.78
Investment contract funds	52	7.53	14.52
Money market funds	5	0.72	7.25
Unfunded funds	3	0.43	0.00
Equity Funds	*548*	*79.31*	*56.00*
Aggressive growth funds	58	8.39	3.08
Company stock funds	193	27.93	15.08
Growth and income funds	66	9.55	24.06
Growth funds	167	24.17	11.33
International equity funds	64	9.26	2.46
Total	691	100.00	100.00

Fund number is the number of funds ever selected as an investment option.

Table 3: Plan Features

	Mean	Std	Min	Max
Offer Investment Menu				
Number of funds	16.59	12.93	1.00	86.00
Balanced Funds	2.99	2.29	0.00	11.00
Equity Funds[1]	9.84	8.12	0.00	59.00
Fixed Income Funds[2]	3.77	3.52	0.00	19.00
Offers Life-Cycle Funds (yes=1)	48.8%			
Offers International Equity Funds (1=yes)	93.1%			
Offers Index Equity Fund (1=yes)	98.8%			
Offers Company Stock (1=yes)	15.0%			
Brokerage Option (1=yes)				
Loans available	74.3%			
Plan Size				
Total Assets per Plan (Million $)	38	146	0	3,624
Active Account Number per plan	776	2,503	1	59,350

1.Observation number: 1,530 plans.

2. Equity Funds=Equity Funds+International, and company stocks are included in Equity Funds.

3. Fixed Income Funds=Bond+Money Market Fund (MMF)+Investment Contracts (IC) +other stable value funds.

4. As of January 2003.

Table 4: Descriptive Statistics of 401(k) Account Holders

	Mean	Std	Min	Max
Panel A: Demographic Features				
Age (years)	43.48	9.90	16.00	83.00
Gender				
Male	47.5%			
Female	26.7%			
Missing	25.8%			
Plan Tenure (years)	7.95	6.76	0.00	56.84
Household Income ('000$)	88	60	8	254
IXI Wealth ('000$)	75	196	-52	2,172
Panel B: Account Features				
# of Selected Funds	3.48	2.22	1.00	49.00
Ratio of Employee Contribution [1]	57.5%	30.4%	0.0%	100.0%
Web Access (1=Yes, 0=No)	37.4%			
Account Balance ('000 $)	86	136	0	17,825

1. Ratio of the balance attributable to employee contribution to the total balance
2. Sample size=1,186,554.

Table 5: Plan-level Trading Activity

	Trading Propensity[1]		Trade Number		Turnover Ratio[2]		# of Plans	# of Participants
	Mean (%)	Std (%)	Mean	Std	Mean (%)	Std (%)		
All	20.51	40.38	0.60	2.67	89.68	236.13	1,530	1,186,554
Number of Funds Offered								
1-5	14.39	23.91	0.40	2.03	62.89	83.97	17	3,963
6-10	17.68	11.35	0.50	2.29	90.08	225.47	478	282,776
11-17	21.48	13.65	0.62	2.78	95.16	273.13	675	572,404
18-30	23.61	13.13	0.55	2.36	79.20	161.71	221	207,918
Over 30	19.63	9.91	0.82	3.40	81.62	172.11	139	119,493
Offers Life-cycle Funds								
No	20.36	13.40	0.60	2.58	86.15	216.31	783	631,769
Yes	20.35	12.39	0.60	2.77	93.79	257.22	747	554,785
Offers International Equity Funds								
No	17.10	14.05	0.43	1.97	77.73	202.22	105	24,711
Yes	20.59	12.80	0.60	2.68	89.88	236.65	1,425	1,161,843
Offers Indexed Equity Funds								
No	51.13	49.99	0.72	1.36	106.53	67.72	19	7,311
Yes	20.32	40.24	0.60	2.68	89.42	237.81	1,511	1,179,243
Loan Available								
No	18.86	15.57	0.45	2.10	86.10	162.56	393	181,023
Yes	20.87	11.82	0.63	2.76	90.24	245.54	1,137	1,005,531
Offers Company Stock								
No	19.77	13.01	0.46	2.10	87.48	190.56	1,301	565,340
Yes	23.66	11.88	0.72	3.10	91.28	264.35	229	621,214
Offers Brokerage Option								
No	19.91	39.93	0.56	2.56	90.24	238.41	1,483	1,132,716
Yes	33.30	47.13	1.32	4.35	82.67	205.20	47	53,838

1.sample trader number/active participant number

2.condititional on trading.

Table 6: Participant-level Trading Activities

	Trading Probability		Trade Number		Turnover Ratio[1]		# of Accounts	# of Traders
	Mean (%)	Std (%)	Mean	Std	Mean (%)	Std (%)		
All	20.51	40.38	0.60	2.67	89.68	236.13	1,186,554	243,398
Age								
Under 35 years old	16.53	37.15	0.40	1.81	78.33	194.86	245,130	40,531
35-44 years old	18.94	39.19	0.53	2.48	88.59	240.42	378,687	71,740
45-54 years old	22.77	41.93	0.72	3.11	94.03	252.38	395,125	89,966
55-64 years old	25.06	43.34	0.78	3.04	93.80	232.45	155,869	39,061
Over 65 years old	17.88	38.32	0.47	1.82	83.15	132.75	11,743	2,100
Plan Tenure								
Under 5 years	17.39	37.90	0.45	2.11	85.13	196.69	522,620	90,877
5-14 years	20.76	40.56	0.59	2.58	89.37	226.06	488,796	101,477
15-24 years	28.59	45.18	1.03	3.89	95.94	291.11	143,730	41,086
Over 25 years	31.71	46.53	1.25	4.66	108.58	372.44	31,408	9,958
Sex								
Male	24.07	42.75	0.75	3.06	96.09	263.63	563,647	135,673
Female	16.09	36.74	0.36	1.74	70.41	150.51	316,821	50,980
Missing	18.54	38.86	0.56	2.69	91.67	228.49	306,086	56,745
Household Income[2]								
Low Group	17.43	37.94	0.48	2.36	91.75	260.04	538,829	93,914
Highgroup	23.08	42.13	0.70	2.90	88.38	219.77	647,725	149,484
IXI Wealth[3]								
Low Group	17.04	37.60	0.48	2.31	94.43	252.92	592,204	100,904
Highgroup	23.97	42.69	0.72	2.99	86.32	223.42	594,350	142,494
On-line available								
No	11.55	31.96	0.22	1.28	78.14	190.61	743,114	85,804
Yes	35.54	47.86	1.23	3.97	95.97	257.34	443,440	157,594

1. Conditional on trading.

2. If the household income is less than the median of the sample, then define it as in low group other wise define it as high group, and the median household income is $86,319.

3. If the IXI wealth is less than the median of the sample, then define it as low group, otherwise, define it as in high group, and the median is $35,735.

4. Statistics is for ratio of traders over active accounts per plan.

Table 7: Distribution of Number of Trades per Account

Number of Trades	Percent	Number of Accounts
0	79.5%	943,156
1	10.9%	129,504
2-5	7.4%	87,864
6-50	2.2%	25,585
51-100	0.0%	385
Over 100	0.0%	60

Note: exclude plan-wide trades.

Table 8: Theoretical Preditors

Theoretical Predictor	Trading Propensity	Trade number	Turnover Ratio[1]
Transaction Cost			
Age	+	+	+
Plan tenure	+	+	+
Household income	+	+	+
Wealth	+	+	+
Offering Life-cycle funds	-	-	-
Offering Loan Availibility	-	-	-
On-line Availibility	+	+	?
Overconfident Theroy			
Male	+	+	+
Diversification			
Number of funds offered	+	+	?
Square of number of funds offered	-	-	?
Change of number of funds offered	+	+	?
Offering company stock	+	+	+
Offering international funds	+	+	+
Offering brokerage option	+	+	+
Efficient investment menu			
Indexed equity fund	-	-	-
Overload information and financial literacy			
Number of funds offered	-	-	?
Square of number of funds offered	+	+	?
Change of number of funds offered	-	-	?
Male[2]	+	+	+
Household income[3]	+	+	+

1: conditional on trader.

2,3: as proxies of financial literacy.

Table 9: Multivariate Robust Analysis of the Probability of Trading

Dependent Variable: Trader (1=trader, 0=non-trader, mean=20.5%)							
Label	Mean	Demographics Only			Full Model		
		Estimate	t-value	Marginal	Estimate	t-value	Marginal
Demographic Variables							
Age	43.48	0.004	24.68	0.0010	0.008	50.35	0.0021
Plan tenure	7.95	0.016	79.56	0.0045	0.011	52.43	0.0030
Male (1=yes, 0=no)	0.48	0.263	80.31	0.0739	0.177	49.25	0.0465
Log(household income)	11.17	0.175	85.20	0.0490	0.112	51.35	0.0293
IXI wealth / household income	0.92	0.011	11.01	0.0032	0.009	12.82	0.0023
Other Account Variables							
Web available (1=yes, 0=no)	0.37				0.798	280.56	0.2264
Employee contribution ratio	0.57				0.207	42.430	0.0543
Plan Design Variables							
Number of funds offered	17.72				0.00214	3.57	0.000561
Number of funds squared	495.11				-0.00002	-2.59	-0.000005
Number of funds changed	0.65				0.001	2.06	0.0002
Offers life-cycle funds	0.47				-0.030	-8.54	-0.0079
Offers indexed equity funds	0.99				-1.215	-66.38	-0.4369
Offers international funds	0.98				0.152	13.18	0.0371
Loan available (1=yes)	0.85				-0.083	-14.61	-0.0222
Offers complany stock	0.52				0.100	27.99	0.0262
Offers brokerage option	0.05				0.181	24.98	0.0509
Log(plan assets)	18.75				-0.011	-10.15	-0.00277
intercept		-3.2437	-137.78		-1.8392	-52.44	
Obs		1,186,554			1,186,554		
Concordant Percentage		60.7%			72.7%		
Pseudo R^2		2.5%			10.7%		
-2log(L)		1,174,334			1,075,240		
Model Comparison (Δ2log(L) vs Criteria $\chi2$)		99,094>12.34					

Note: Missing dummies included for all explanotary variables. We also control on industry dummies.
All variables are as of Jan 2003.
Standard errors are cluster-corrected robust.
All independent varialbes are significant at 5% level.

Table 10: Multivariate Robust Analysis of the Number of Trades

Dependent Variable: Total trade number (mean=0.60)					
Label	Mean	Only Demographic Model		Full Model	
		Estimate	t-value	Estimate	t-value
Demographic Variables					
Age	43.48	0.009	9.86	0.016	14.75
Plan tenure	7.95	0.036	26.61	0.024	14.41
Male (1=yes, 0=no)	0.48	0.689	31.26	0.472	17.43
Log(household income)	11.17	0.280	19.17	0.161	9.31
IXI wealth / household income	0.92	0.036	22.80	0.021	12.45
Other Account Variables					
Web available (1=yes, 0=no)	0.37			1.638	75.51
Employee contribution ratio	0.57			0.332	9.30
Plan Design Variables					
Number of funds offered	17.72			0.01099	2.64
Number of funds squared	495.11			-0.00009	-1.62
Number of funds changed	0.65			0.006	2.14
Offers life-cycle funds	0.47			-0.031	-1.24
Offers indexed equity funds	0.99			-1.328	-20.52
Offers international funds	0.98			0.173	2.05
Loan available (1=yes)	0.85			-0.117	-2.84
Offers complany stock	0.52			0.230	8.78
Offers brokerage option	0.05			0.309	6.02
Log(plan assets)	18.75			0.000	-0.02
Scale		6.878		4.787	
intercept		-4.907	-29.62	-3.688	-15.40
Obs		1,186,554		1,186,554	
Pseudo R^2		1.5%		6.6%	
-2log(L)		2,018,282		1,914,053	
Model Comparison (Δ2log(L) vs Criteria χ2)		104,229>12.34			

Note: Missing dummies included for all explanotary variables. We also control on industry dummies.
All variables are as of Jan 2003.
Standard errors are cluster-corrected robust.
Shadow area indicates not significant with 95% confident interval.

Table 11: Second Stage in Two-stage Multivariate Robust Analysis of the Turnover Ratio

Dependent Variable: Turnover Ratio [1] (mean=89.3%)				
	Demographics Only		Full Model	
	estimate	t-value	estimate	t-value
Demographic Variables				
Age	0.0038	7.81	0.0096	6.14
Plan tenure	0.0023	2.53	0.0092	4.11
Male (1=yes, 0=no)	0.2036	20.68	0.3149	9.48
Log(household income)	-0.0832	-10.10	0.0036	0.16
IXI wealth / household income	-0.0069	-6.50	-0.0003	-0.18
Other Account Variables				
Web available (1=yes, 0=no)			0.6535	4.43
Employee contribution ratio			-0.0102	-0.24
Plan Design Variables				
Number of funds offered			-0.0073	-4.72
Number of funds squared			0.0001	4.11
Number of funds changed			0.0076	5.46
Offers life-cycle funds			0.1036	8.29
Offers indexed equity funds			-0.7247	-3.36
Offers international funds			0.1517	3.36
Loan available (1=yes)			-0.0054	-0.24
Offers complany stock			-0.0161	-0.72
Offers brokerage option			0.0721	1.86
Log(plan assets)			0.0322	8.02
INVERSE MILLS RATIO	-0.2461	-18.11	0.7548	3.16
intercept	1.7961	18.09	-1.1277	-2.10
Obs	243,398		243,398	
Pseudo R^2	0.02%		0.1%	
-2log(L)	1,108,154		1,107,619	
Model Comparison (Δ2log(L) vs Criteria $\chi 2$)	535>12.34			

1.Conditional on traders.

Note: Missing dummies included for all explanotary variables. We also control on industry dummies.
All variables are as of Jan 2003.
Standard errors are cluster-corrected robust.
Shadow area indicates not significant with 95% confident interval.

Table 12: Summary of the Marginal Effects

	Probability of Trading[1] (mean=20.5%)	Number of Trades[2] (mean=0.60)	Turnover Ratio[3] (mean=89.3%)
Demographic Variables			
Age	0.21%	1.63%	0.96%
Plan tenure	0.30%	2.38%	0.92%
Male (1=yes, 0=no)	4.65%	47.24%	31.49%
Log(household income)	2.93%	16.14%	0.36%
IXI wealth / household income	0.23%	2.09%	-0.03%
Other Account Variables			
Web available (1=yes, 0=no)	22.64%	163.76%	65.35%
Employee contribution ratio	5.43%	33.18%	-1.02%
Plan Design Variables			
Number of funds offered	0.06%	1.10%	-0.73%
Number of funds squared	-0.001%	-0.01%	0.01%
Number of funds changed	0.02%	0.64%	0.76%
Offers life-cycle funds	-0.79%	-3.14%	10.36%
Offers indexed equity funds	-43.69%	-132.83%	-72.47%
Offers international funds	3.71%	17.25%	15.17%
Loan available (1=yes)	-2.22%	-11.65%	-0.54%
Offers complany stock	2.62%	23.03%	-1.61%
Offers brokerage option	5.09%	30.86%	7.21%

1.Marginal effects are computed at the mean values of the independent variables.
2.Marginal effects indicate percentage change of the number of trades caused by one unit change in independent variables
3.Conditional on traders.
Shadow area indicates not significant with 95% confident interval.

Appendix 1: Plan Status Identification

We identified participants' plan status with their contribution records. The basic idea is if an employee (EE) or his/her employer (ER) is contributing to his/her 401(k) account in a given month, then we define his/her plan status as active. If an employee is active for the whole sample period, then we define him/her as sample active, and we focus the observations with sample active plan status.

Because it is impossible to identify the ACCURATE contribution schedule for each account, we identified the plan status with the following decision rule.

- Step 1: Categorize plans into 3 types by plan-level sample total contribution amount.
  - If the plan-level total employee-contribution amount in sampled 2 years is equal to zero, and plan-level total employer-contribution amount is greater than zero, then define this plan as 'ER only' type.
  - If the plan-level total employer-contribution amount in sampled 2 years is equal to zero, and plan-level total employee-contribution amount is greater than zero, then define this plan as 'EE only' type.
  - If both the plan-level total employer-contribution amount and plan-level total employee-contribution amount in sampled 2 years are greater than zero, then define this plan as 'EE and ER' type.
- Step 2: Define an account as "**Active in a given month**" according to the following rule:
  - **<u>Case 1</u>:** The plan is an ER-only contribution plan.

    →RULE: A given <u>account</u> is classified as *active in that month* if that account has <u>any</u> ER contribution within a plus/minus 12-month window, otherwise defines it as inactive.
  - **<u>Case 2</u>**: The plan is an EE-only contribution plan.

    →RULE: An account is *active that month* if the account has a positive EE contribution in that month, otherwise define it as inactive.
  - **<u>Case 3</u>:** The plan has both EE and ER contributions.

    → In this case, an account is *active that month* if the account has positive EE contribution;

→ELSE, for each <u>account,</u> compute EE contribution frequency over all months to ascertain whether the EE contribution has been positive at any time in the last 6 months.

→THEN consider 2 cases as follows

- *An account* is active there is no any EE contribution in all months, <u>but</u> has some positive ER contribution within plus/minus 12-month window, otherwise define it as inactive.
- An account is active if there is some positive EE contribution in any of last 6 months, otherwise define it as inactive.

- Step 3: Define an account to be "Continuously Active" if and only if it is active in all of the months from Jan 2003-Dec 2004. This is the group we would focus on in our trading analysis.

Appendix 2: Plan-wide Trading

In our sampled observations, we excluded plan-wide trades from our analysis. The reason is that plan-wide trades are not based on individual investors' own decision-making, but because employers drop some funds from the investment menu, employees are "forced" to exchange all of the balance in those funds to other possible options.

In practice, we defined plan-wide trades, only-plan-wide-trader as below.

If a fund satisfies either of the following two conditions in a given month, then we define <u>*all*</u> of the trades associated with that funds in <u>*that month*</u> as plan-wide trades. The conditions are

1) The fund balance in a given plan is less than 2% of that in previous month.
2) Some funds are traded in a given month by some accounts, but in that month, there is no balance even in plan to which those accounts belong.

We also defined some other trading concepts associated with plan-wide trades, and treat plan-wide trade as below in order to generate our interest dependent variables.

1) Only-plan-wide-trader: consider a sample trader as Only-plan-wide-trader if all of his/her trades within sampled 2 years are plan-wide trades. We regard Only-plan-wide-trader as non-trader.

2) Sample Trade Number: we excluded plan-wide trades, e.g, if a participants trade 10 times, of which 3 of them are plan-wide trades during sample period, then his sample trade number is 7.
3) Sample Turnover Ratio: we didn't count plan-wide trading amount as part of the total trading amount, e.g, if a participant trades twice with amount of $70 and 50$, and the $50 trade is a plan-wide trade, given his average balance as $100, then his sample turnover ratio is 70% ($70/$100=70%).

References

Agnew, Julie and Szykman Lisa R., "Asset Allocation and Information Overload: the Influence of Information Displan, Asset Choice and Investor Experience", *CRR Working Paper* 2004-15, April 2004

Agnew, Julie, Pierluigi Balduzzi, and Annika Sunden, "Portfolio Choice and Trading in a Large 401(k) Plan", *American Economic Review*, March 2003 pp193-215

Amemiya, Takeshi *Advanced Econometrics* Chapter9 (Harvard Press 1985)

Ameriks, John and Zelders, Stephen P., "How Do Household Portfolio Shares Vary With Age?" *Working Paper*, September 2004

Barber, Brad M. and Terrance Odean, "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment", *Quarterly Journal of Economics*, February 2001, Vol. 116, No. 1, 261-292.

Barber, Brad M. and Terrance Odean, "The Internet and the Investor", *Journal of Economic Perspectives*, Vol.15, Nov 1-Winter, 41-54.

Barber, Brad M. and Terrance Odean, "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors", *Journal of Finance*, Vol. LV, No. 2, April 2000, 773-806.

Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean. 2004. "Who Gains from Trade? Evidence from Taiwan." Working paper. faculty.haas.berkeley.edu/odean/Current%20Research.htm.

Benartzi, Shlomo and Richard H. Thaler. 2001. "Naive Diversification in Defined Contribution Savings Plans." *American Economic Review.* 91(1). 79-98.

Bernheim, B. Douglas. 1998. "Financial Illiteracy, Education and Retirement Saving." In Olivia S. Mitchell and Sylvester J. Schieber, eds., *Living with Defined Contribution Plans.* University of Pennsylvania Press, Philadelphia, PA. 38-68.

Bronson, James W., Matthew H. Scanlan and Jan R. Squires. Forthcoming. "Managing Individual Investor Portfolios." In *Managing Investment Portfolios: A Dynamic Process,* 3rd Edition. CFA Institute, Charlottesville, VA.

Burton T. Beam, Jr. John J. McFadden, *Employee Benefits*, Limited Edition not for release, 2004

Choi, James J., David Laibson, and Andrew Metrick, "How does the Internet affect trading? Evidence from investor behavior in 401(k) plans", *Journal of Financial economics* 64(2002) 397-421

Choi, James J., David Laibson, and Brigitte C. Madrian. 2004. "Plan Design and 401(k) Saving Outcomes." *National Tax Journal.* LVII(2). 275-298.

Choi, James J., David Laibson, and Madrian, Brigitte, "Employer Stock Holding in 401(k) Plans: Have Employees Learned the Lessons of Enron?", *Working Paper* July 2004

Choi, James J., David Laibson, Brigitte C. Madrian and Andrew Metrick. 2002a. "Defined Contribution Pensions: Plan Rules, Participant Choices and the Path of Least Resistance." In James Poterba, ed., *Tax Policy and the Economy*. MIT Press, Cambridge, Massachusetts. 16. 67-113.

Choi, James J., David Laibson, Madrian, Brigitte and Andrew Metrick, "For Better or For Worse: Default Effects and 401(k) Savings Behavior",*NBER Working Paper* Nov 2001

DeRoon, Nijman and Werker, "Testing for mean variance spanning with short sales constraints and transaction costs: the case of emerging market." *Journal of Finance* 56(2), 721-742

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake. 2004. "The Adequacy of Investment Choices Offered by 401(k) Plans." Stern School Working Paper, New York University.

Even, William E. and David A. Macpherson. 2004. "Determinants and Effects of Employer Matching Contributions in 401(k) Plans." Working Paper. www.sba.muohio.edu/evenwe/papers/matching_in_401k.pdf.

Gervais, Simon and Terrance Odean, "Learning to be Overconfident", *Review of Financial Studies*, Spring 2001, Vol. 14, No. 1, pp. 1-27.

Goetzmann, William N., Massimo Massa, and K. Geert Rouwenhorst, "Behavioral Factors in Mutual Fund Flows" International Center for Finance at the Yale School of Management *Working paper*.

Greene, William, *Econometrics Analysis*, Chapter 17, 21-22, Prentice Hall (2003)

Grinblatt, Mark and Matti Keloharju, "What Makes Investors Trade?", *J of Finance*, April 2001

Grossman, Sanford. J and Joseph E. Stiglitz, "On the Impossibility of Informational Efficient market", *American Economic Review*, Vol.70, No.3 (June,1980) p393-408

Holden, Sarah and Jack VanDerhei. 2004. "401(k) Plan Asset Allocation, Account Balances and Loan Activity in 2003." *ICI Perspective.* Investment Company Institute, Washington DC. 10(2).

Iyengar, Sheena, Gur Huberman, and Wei Jiang. 2004. "How Much Choice is Too Much? Contributions to 401(k) Retirement Plans." In *Pension Design and Structure: New Lessons from Behavioral Finance.* Olivia S. Mitchell and Stephen P. Utkus, eds. Oxford, United Kingdom: Oxford University Press: 83-95.

King, Mervyn A., and Leape Jonathan I., "Asset Accumulation, Information, and the Life Cycle" *NBER Working Paper* No.2392, September 1987.

Madrian, Brigitte, and D.F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116: 1149-1525.

McCullagh, P and Nelder, J.A. *"Generalized Linear Models"* (Chamman & Hall/CRC 1989)

Mitchell, Olivia S. and Stephen P. Utkus. 2004. "The Role of Company Stock in Defined Contribution Plans." In Olivia S. Mitchell and Kent Smetters, eds., *The Pension Challenge: Risk Transfers and Retirement Income Security*. Oxford University Press. 33-70.

Mitchell, Olivia S., Stephen P. Utkus, and Tongxuan (Stella) Yang. 2005. “Turning Workers into Savers? Incentives, Liquidity and Choice in 401(k) Plan Design.” NBER Working Paper 11726.

Mitchell, Olivia, Alexander Muermann, and Jacqueline M. Volkman, “ Regret, Portfolio Allocation, and Pension Guarantees”, Department of Insurance and Risk Management, University of Pennsylvania, *Working Paper 2004*

Odean, Terrance, "Do Investors Trade Too Much?", *American Economic Review*, Vol. 89, December 1999, 1279-1298.

Odean, Terrance, "Volume, Volatility, Price, and Profit When All Traders Are Above Average," *Journal of Finance*, Vol. LIII, No. 6, December 1998, 1887-1934.

Samuelson, Paul A. 1969. “Lifetime Portfolio Selection by Dynamic Stochastic Programmin g.” *Review of Economics and Statistics.* 51(3). 239-246.

Stuart A. Klugman, Harry H. Panjer, and Gordon E. Willmot, *“Loss Model: from Data to Decisions”* 2nd Edition, Wiley (2004)

Tacchino, Kenn Beam and David A. Littel. *Planning for Retirement Needs.* Third edition. The American College, Bryn Mawr, PA. 1999.

Vanguard. 2004. *How America Saves: A Report on Vanguard Defined Contribution Plans 2004.* Vanguard Center for Retirement Research, Malvern, PA. www.vanguardretirementresearch.com.