Dynamic Stock Market Participation of Households
with Heterogeneous Participation Costs

Natalia Khorunzhina

Department of Economics
Copenhagen Business School
Porcelnshaven 16 A, 1
2000 Frederiksberg, Denmark
Phone: +45 3815 2403
Fax: +45 3815 2576
nk.eco@cbs.dk

June, 2012
Abstract

This paper develops and estimates a dynamic model of stock market participation, where consumers’ decisions regarding stock market participation are influenced by participation costs. The practical significance of the participation costs is considered as being a channel through which financial education programs can affect consumers’ investment decisions. Using household data from the Panel Study of Income Dynamics, I estimate the magnitude of the participation cost, allowing for individual heterogeneity in it. The results show the average stock market participation cost is about 5% of labor income; however, it varies substantially over consumers’ life. The model successfully predicts the level of the observed participation rate and the increasing pattern of stock market participation over the consumers’ life cycle.

JEL D12, D14, D91

Keywords: Stock Market Participation, Dynamic Models, Discrete Choice Models.
1 Introduction

Despite the theoretical prediction that all investors will participate in stock markets if the equity premium is positive, empirical evidence shows a substantial number of consumers do not invest in stock markets either directly or indirectly (via pension account schemes, mutual funds, or similar institutions)[1] Over the past decade, the so-called limited stock market participation phenomenon has received growing attention in both the theoretical and empirical literature. One of the popular explanations for the phenomenon is the existence of stock market participation costs that arise from the time and effort necessary for obtaining and processing financial knowledge and information, for following the current trends on financial markets, for paying sign-up fees, for filing the necessary paperwork associated with stock holdings, and so on. Stock market participation costs, however, are not observed by researchers, which comprises the major difficulty in quantifying it. Yet using the data to reveal not only the magnitude of the participation cost but also its sources of heterogeneity is important for a number of reasons. First, participation cost is one of the parameters in life-cycle models of portfolio allocation and wealth accumulation; therefore, its various magnitudes can lead to different implications and may result in different economic policy conclusions. Second, economic programs that aim to provide financial education to consumers may lead to reduced participation costs, thereby encouraging stock market participation. However, the effect of such programs is likely to differ for different groups of consumers. This paper develops and estimates empirically a dynamic model of stock market participation, in which heterogeneous participation costs influence consumers’ decisions regarding stock market participation. Using household data from the Panel Study of Income Dynamics, I estimate the magnitude of the stock market participation cost, allowing for individual heterogeneity in participation costs, represented by age, education, and participation experience.

Over the past century, average returns on equity have far exceeded the average returns on risk-free assets. In light of impressive equity premium over these years, the unwillingness to invest in stocks is arguably an investment mistake (Campbell, 2006; Calvet, Campbell, and Sodini, 2007). The launching of financial education curricula via a number of economic policy incentives a decade ago acknowledged the importance of the promotion of consumer financial education[2] Further, a growing concern in the literature and in economic policy debates reinforced the issue of limited stock market participation (see Guiso and

---


[2] Examples are the Economic Growth and Tax Relief Reconciliation Act, and Money Smart, a program launched by the Federal Deposit Insurance Corporation.
Recently, the debates resulted in a number of financial education programs designed to further increase and promote financial literacy among consumers. Financial education and counseling alleviates the burden on consumers’ time and the effort necessary for making financial decisions and reduces the objective cost of stock market participation. In life-cycle models of portfolio allocation and wealth accumulation, participation cost may be viewed as a modeling tool that allows us to investigate whether financial education programs can influence consumers’ financial choices and increase participation in financial markets.

Although stock market participation costs are not observed, the literature has collected supporting evidence that participation costs are non-negligible. Theoretical simulations of Haliassos and Michaelides (2003) and Gomes and Michaelides (2003), followed by empirical studies of Vissing-Jorgensson (2002), Paiella (2007), and Attanasio and Paiella (2011) show that a stock market entry cost or participation cost can rationalize not participating in stock markets. However, the literature finds little agreement on the size of the participation costs. Alan (2006) finds stock market entry costs equal about 2% of annual permanent income, while Haliassos and Michaelides (2003) obtain a wide range for entry costs, from 3% to 34% of mean annual income. The existing empirical estimates of per-period participation costs in Attanasio and Paiella (2011) and Paiella (2007) provide only a lower bound in units of nondurable consumption as low as 0.4% and 0.7% of consumption per year, arguing that the true costs of participation may exceed these levels in reality. Although the empirical findings of Attanasio and Paiella (2011) and Paiella (2007) show investor behavior is sensitive even to small changes in consumption, the evidence on lower bounds of participation costs give limited information for further analysis of economic welfare. A sharper estimate of participation costs can be beneficial for evaluation of policy interventions in models of lifetime consumption and portfolio choice. One of the novelty of the paper is that it estimates empirically the magnitudes of the participation costs. Then I go one step further by exploring the heterogeneity of participation costs, which constitutes another novelty of the paper.

The empirical studies cited earlier necessarily adopt a broad definition of the participation costs, which include both pecuniary (financial) and non-pecuniary (behavioral) components, because insufficient data usually hinders separate identification of these components. The notable exception is the work of Andersen and Nielsen (2011), whose unique data allow for separation of the effects of financial and behavioral costs.

---

3 The latest incentive is the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010.

4 These estimates translate into a $72 bound for Attanasio and Paiella (2011) and a $130 bound for Paiella (2007). Vissing-Jorgensen (2002) finds the bound of per-period fixed costs as low as $260 can explain the behavior of most nonparticipants.
They find limited stock market participation is not likely to be driven by financial constraints but rather by behavioral, cognitive, and psychological barriers. Duflo and Saez (2000) and Hong, Kubik, and Stein (2004) investigate how one’s peers affect participation, offering additional evidence in favor of behavioral motives and cognitive abilities as barriers to stock market participation. Grinblatt, Keloharju and Linnainmaa (2009) show a link between IQ and stock market participation. Although the structural model developed in this paper does not directly connect to this strand of the literature, the flexible form of participation costs enables me to account for both financial and non-financial barriers to stock market participation and evaluate the cohort effects in participation. In doing so, I allow the participation costs to depend on investor’s education as a proxy for the ability to collect and process information and on age and past participation as proxies for the accumulation of information and experience.

My analysis departs from the life-cycle consumption and portfolio choice model, as in Gomes and Michaelides (2003), where households optimize lifetime utility subject to the wealth-accumulation budget constraint. The key component of the analysis of participation cost is the intertemporal utility cost and benefit that arises with different participation strategies. Stock market participation costs have a direct effect on savings. They are also connected to the share of wealth households choose to invest in stocks. On the one hand, the participation cost reduces savings today - a risk a household takes in the hopes of receiving a greater expected return on its portfolio and a corresponding increase in wealth in the next period. On the other hand, households may choose not to invest in stocks today, but instead keep all savings in risk-free assets. In the next period, however, the household receives a risk-free return on wealth, which also includes the not-foregone participation cost. The estimation of the participation costs relies on comparison of the value functions associated with the two participation strategies.

The estimation results provide evidence that the participation cost, measured as a share of income, can be substantial. It is indeed, on average, greater than the lower-bound estimates of Paiella (2007) and Attanasio and Paiella (2011). The average stock market participation cost is estimated to be about 5% of labor income; however, it varies dramatically over the consumer’s life. The results show the participation costs are less for more educated investors and decrease as consumers become older. When age enters the participation cost in a quadratic form, the participation cost is increasing and concave in age, with the turning point at age 35-40. I find the evidence that the participation costs are substantially smaller for households who continue to be stock market participants. The results suggest that the per-period cost decreases by more than 3% of income when households participate in previous periods. The estimated model also implies
the stock market participation costs are larger for first-time investors. Therefore, the estimation results on stock market participation costs not only support the previous findings in empirical literature, but also are in accordance with the fixed entry costs used in theoretical literature. The striking difference of the estimated structural model with similar models in theoretical studies lies, however, in successful prediction of the low stock market participation rate observed in the data.

The rest of the paper is organized as follows. Section 2 describes data sample construction. Section 3 outlines the model. Section 4 presents the empirical implementation. Section 5 contains the empirical findings. Section 6 illustrates how the model fits the data and provides robustness checks. Section 7 concludes.

2 Data

Wealth supplement of the the Panel Study of Income Dynamics (PSID) is a longitudinal data that covers all age cohorts, which makes it suitable for studying dynamic stock market participation of prime-age consumers. Starting in 1999, the wealth supplement of the PSID is available every two years. In my analysis, I use observations on household liquid wealth, income, and demographic characteristics for five time periods, from 1999 through 2007.

Household stock market participation is a result of a complicated decision making that involves taking into consideration various background risks and other investment opportunities. Risky labor income may cause a crowding out of stock market investments (Benzoni, Collin-Dufresne, and Goldstein (2007)). Heaton and Lucas (2000) argue private business assets substitute for public equity. Homeownership and changes in housing may result in shifts of household portfolio away from direct stock holdings (Grossman and Laroque, 1990; Flavin and Yamashita, 2002; Chetty and Szeidl, 2007). Finally, deep changes in family composition affect household financial portfolio as well (Love, 2010). In order to keep focus on stock market participation in the current study, I try to minimize the potential influence of these factors, and consider households that satisfy the following conditions: (i) do not invest in business and/or farm; (ii) do not experience changes in housing; and (iii) have the same head of household for the whole observation period. Table 1 shows summary statistics for the data I use in my analysis.

The value of households’ stock holdings is observed in the PSID as “non-IRA stock holdings”.

---

5 Other high quality wealth data used in household finance are the US Survey of Consumer Finances (SCF), and the US Health and Retirement Survey (HRS); however, the use of these data is not appropriate for the current study due to the cross-section dimension of SCF and the emphasis on older cohorts in HRS. For the alternatives beyond the aforementioned US data sets, see the review on comparability of the existing micro-data on household wealth in Sierminska, Brandolini, and Smeeding (2006).

6 The PSID separately provides information on the total amount held in individual retirement accounts and an approximate split
liquid wealth is computed as a sum of non-IRA stock holdings, money in checking or savings accounts, money market funds, certificates of deposit, government savings bonds, Treasury bills, other bond holdings, and money in other savings or assets, such as bond funds. The value of liquid wealth that is not invested in stocks is considered a risk-free asset. Total household income consists of labor income, financial income (interest, dividends, income from trust fund), and other money inflow (child support, help from relatives, rent, worker’s compensation) of the head of the household and the spouse. I remove extreme outliers by excluding observations for wealth and income above the 99th percentile and below the 1st percentile.* Data on income and wealth are deflated using the CPI for the end of the year before the interview was conducted. The CPI deflator is taken from the consumer price index releases of the Bureau of Labor Statistics.

Table 1 shows the dramatic difference in both income and wealth between participants and nonparticipants. On average, nonparticipants have a lower income. The accumulated wealth of nonparticipants is substantially lower as well. However, larger standard deviations accompany the greater average income and wealth of stock market participants.

Although data on wealth and income are available, the PSID does not report total household consumption. To circumvent this problem, I compute the measure of a household’s consumption from the budget constraint. A comparable imputation has been previously done with the PSID data (see Ziliak, 1998). Other studies that methodologically justify the imputation of consumption expenditure based on the information on income and wealth include Browning and Leth-Petersen (2003) and Koijen, Nieuwerburgh and Vestman (2011). The latter study provides further evidence that the imputation of consumption from the budget constraint may overcome measurement deficiencies present in survey-based consumption, especially for the wealthy. The precision of the imputed consumption necessarily depends on the quality of income and wealth data. The studies of Browning and Leth-Petersen (2003) and Koijen, Nieuwerburgh and Vestman (2011) take advantage of the indisputable accuracy of the administrative data to produce a reliable measure of consumption. Due to the shortage of administrative data on wealth for the US households, the accuracy of the imputed consumption relies on the quality of survey-based wealth data. The inquiry into wealth data of Bosworth and Smart (2009) suggests an equal grade of wealth data in the PSID and other credible surveys, the US HRS and the US SCF, with the divergence for only top five percent of households. It is mentioned above that the extremely wealthy consumers are not represented in the data sample used in the analysis, of the amount in an IRA between interest- and dividend-earning assets. The exact value of an IRA in stock holdings is not available.

*The portfolio choice of the rich is not in the scope of the current study, however, it deserves a special attention as risky asset allocation for the wealthy is in striking difference to the one for the remaining population (see Carroll, 2002).
### Table 1: Summary Statistics

<table>
<thead>
<tr>
<th>Year</th>
<th>1999</th>
<th>2001</th>
<th>2003</th>
<th>2005</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observations</strong></td>
<td>1,332</td>
<td>1,424</td>
<td>1,382</td>
<td>1,481</td>
<td>1,524</td>
</tr>
<tr>
<td><strong>Income and Wealth</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>39,420.6</td>
<td>39,532.9</td>
<td>38,748.0</td>
<td>40,040.0</td>
<td>40,252.8</td>
</tr>
<tr>
<td>Nonparticipants</td>
<td>34,716.9</td>
<td>35,363.8</td>
<td>35,698.1</td>
<td>36,913.2</td>
<td>36,965.0</td>
</tr>
<tr>
<td>(19,885.0)</td>
<td>(21,087.4)</td>
<td>(21,198.6)</td>
<td>(22,709.3)</td>
<td>(22,290.0)</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>50,740.6</td>
<td>50,243.1</td>
<td>46,938.1</td>
<td>49,565.4</td>
<td>51,530.6</td>
</tr>
<tr>
<td>(29,623.8)</td>
<td>(27,695.0)</td>
<td>(29,134.3)</td>
<td>(30,019.2)</td>
<td>(29,623.2)</td>
<td></td>
</tr>
<tr>
<td>Wealth</td>
<td>21,525.4</td>
<td>21,974.9</td>
<td>23,793.2</td>
<td>22,864.0</td>
<td>21,859.5</td>
</tr>
<tr>
<td>Nonparticipants</td>
<td>10,407.1</td>
<td>10,725.3</td>
<td>11,440.3</td>
<td>10,338.6</td>
<td>10,532.2</td>
</tr>
<tr>
<td>(23,649.3)</td>
<td>(25,346.4)</td>
<td>(27,277.2)</td>
<td>(22,870.9)</td>
<td>(25,020.8)</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>48,283.5</td>
<td>50,874.2</td>
<td>56,965.1</td>
<td>61,022.1</td>
<td>60,714.7</td>
</tr>
<tr>
<td>(58,981.7)</td>
<td>(64,580.2)</td>
<td>(71,231.2)</td>
<td>(74,778.4)</td>
<td>(71,578.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Stock holders</strong></td>
<td>0.29</td>
<td>0.28</td>
<td>0.27</td>
<td>0.25</td>
<td>0.22</td>
</tr>
<tr>
<td><strong>Share of wealth in stock</strong></td>
<td>0.61</td>
<td>0.58</td>
<td>0.53</td>
<td>0.55</td>
<td>0.57</td>
</tr>
<tr>
<td><strong>Stock holdings of participants</strong></td>
<td>32,883.1</td>
<td>32,979.9</td>
<td>32,451.9</td>
<td>37,838.5</td>
<td>37,270.7</td>
</tr>
<tr>
<td><strong>Demographic data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>43.3</td>
<td>44.3</td>
<td>45.1</td>
<td>46.1</td>
<td>46.0</td>
</tr>
<tr>
<td>Nonparticipants</td>
<td>42.6</td>
<td>43.7</td>
<td>44.6</td>
<td>45.4</td>
<td>45.4</td>
</tr>
<tr>
<td>Participants</td>
<td>44.8</td>
<td>46.0</td>
<td>46.5</td>
<td>48.3</td>
<td>48.0</td>
</tr>
<tr>
<td>Family size</td>
<td>3.03</td>
<td>3.00</td>
<td>2.93</td>
<td>2.94</td>
<td>2.89</td>
</tr>
<tr>
<td>Nonparticipants</td>
<td>3.10</td>
<td>3.02</td>
<td>2.97</td>
<td>2.99</td>
<td>2.92</td>
</tr>
<tr>
<td>Participants</td>
<td>2.87</td>
<td>2.94</td>
<td>2.82</td>
<td>2.80</td>
<td>2.78</td>
</tr>
<tr>
<td>Education</td>
<td>13.7</td>
<td>13.6</td>
<td>13.6</td>
<td>13.7</td>
<td>13.7</td>
</tr>
<tr>
<td>Nonparticipants</td>
<td>13.3</td>
<td>13.3</td>
<td>13.3</td>
<td>13.4</td>
<td>13.4</td>
</tr>
<tr>
<td>Participants</td>
<td>14.6</td>
<td>14.6</td>
<td>14.5</td>
<td>14.6</td>
<td>14.7</td>
</tr>
<tr>
<td>Marital status</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>Nonparticipants</td>
<td>0.69</td>
<td>0.67</td>
<td>0.67</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>Participants</td>
<td>0.74</td>
<td>0.78</td>
<td>0.76</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Occupation</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td>n.a.</td>
<td>n.a.</td>
<td>0.16</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td>Nonparticipants</td>
<td>0.13</td>
<td>0.15</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>0.23</td>
<td>0.23</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Standard deviation is reported in parenthesis.

<sup>b</sup>Equals to 1 if related to management, business operation, or financial specialist; 0 otherwise.
therefore the imputed consumption can be considered a reliable measure of the true consumption.

Demographic characteristics include age, education, and occupation of the head of the household, as well as marital status and family size. Individual consumption and savings behavior differs depending on the consumer’s stage of life: in prime working age or at retirement. This study mostly considers prime-age consumers and excludes households whose head is either younger than 22 or older than 65 over the period of the estimation. I also exclude households whose marital status changed over the sample period.

Table 1 shows differences in demographic characteristics of participants and nonparticipants. On average, participants are two to three years older and more educated. Their occupation is typically more related to management, business operations, or finance. More married individuals are among stock market participants compared to nonparticipants. The difference in family composition is not substantial.

Uncertainty about returns on risky assets plays an important role in dynamic decisions about participation and portfolio composition. The distribution of the rate of return on stockholdings is evaluated using the information on the SP500 index for the period. The real risk-free rate $R_f$ is constructed based on seasonally adjusted deflated average six-month Treasury bill.

For the analysis that follows, I take into account the data restrictions discussed above and construct three data samples. The samples include a data subset used in estimation of the preference parameters and the parameters of the participation cost, a data subset with an extended time frame for estimation of the earnings’ equation, and a data subset with an extended number of individuals to estimate the conditional choice probabilities.

The first stage of analysis involves the estimation of the individual effects. I use the earnings equation to estimate individual unobserved fixed effects. To reduce the bias in fixed effects, one needs a data sample with a larger time dimension. Unlike the wealth data, labor income is reported in the PSID in all time periods well beyond the period for which the wealth supplement is available. To estimate the earnings’ equation, I take the extended data set from 1981 through 2007. Because the main model only uses data on odd years, for consistency, I use data on odd years only. Using the age and income restrictions described above, I construct an extended data set of 7,744 households that contains data on each household’s labor income, age, education, and family size.

The second stage of the analysis estimates conditional choice probabilities that also require a large number of observations, especially along the cross-sectional dimension of the panel. This data set is constructed using wealth variables, among others, so it has time dimension from 2001 (accounting for one lag in wealth).
through 2007. To ensure a larger number of observations, I drop the age and marital status restrictions discussed above. Namely, I use households of all ages and disregard the changes in marital status while still controlling for marital status, and age. The conditioning set for estimation of conditional choice probabilities includes past wealth, past share in wealth invested in stocks, return on household portfolio, income, age, education, occupation, family size, and marital status as well as estimated individual effects and time dummies from the earnings’ equation. A total of 10,708 observations are available over four time periods.

Finally, the third stage of the analysis estimates the preference parameters and the parameters of the participation cost. This data sample is constructed as an unbalanced panel of 299 households observed over three periods: 2001, 2003, and 2005. Two other time periods (1999 and 2007) are lost while taking one lag and one lead. I observe these households participating in stock markets in the current period. Regarding the treatment of the self-selection issue, note the following: first, although the utility parameters are estimated with stock market participants, the construction of the estimator allows me to fully use the information on nonparticipants in the estimation of the individual effects and the conditional choice probabilities in the previous stages of estimation; therefore, I incorporate this information into the estimation of the final stage through the estimates of unobserved effects and CCPs. Second, I construct the third-stage estimator from an identity equation that holds equally for participants and nonparticipants.

The set of model covariates includes past and current wealth, the past and current share of wealth invested in stocks, current and future income, current demographic characteristics, as well as estimates of the conditional choice probabilities, and transition probabilities, computed from the empirical distribution of returns on stockholdings. The instrument set for orthogonality conditions includes variables from the state vector: the past share of wealth invested in stock, past portfolio allocation, return on household portfolio, income, and family characteristics.

3 Theoretical Model

\(N\) households are indexed by \(i = 1, \ldots, N\), each with a lifespan of \(T\) periods. In each period, household \(i\) observes its accumulated wealth and a current rate of return on the portfolio composed of riskless and (possibly) risky assets. After observing income, the household observes the amount of cash on hand to be allocated between a single composite consumption good and savings.\(^8\) The household also decides how to reallocate savings between risky and riskless asset. In each period, the household decides to either invest in

\(^8\)I assume household supplies labor in each period.
stocks, denoted by \( d_i = 1 \), or to not invest and instead keep all savings in the riskless assets, in which case \( d_i = 0 \). If the household decides to invest in stocks, it also chooses the share of wealth allocated to stock holdings, represented by \( \alpha_i \).

Let \( z_{it} \) be the state vector of household \( i \) in period \( t \) that is composed of observed and unobserved, but otherwise statistically inferred, characteristics. In every period, household \( i \) receives a utility payoff \( u(z_{it}, d_{it}) + (1 - d_{it})\varepsilon_{0it} + d_{it}\varepsilon_{1it} \), where \( u(z_{it}, d_{it}) \) depends on consumption and stock market participation choices and a pair \((\varepsilon_{0it}, \varepsilon_{1it})'\) is choice-specific utility shocks, independent over \( i \) and \( t \), which is revealed to the household at the beginning of period \( t \). The household’s problem can be formulated as follows:

\[
\max_{d_{it}} \mathbb{E}_t \sum_{t=1}^{T} \beta^{t-t} [u(z_{it}, d_{it}) + (1 - d_{it})\varepsilon_{0it} + d_{it}\varepsilon_{1it}]
\]

subject to:

\[
w_{it} = R^{p}_{it} w_{it-1} + y_{it} - c_{it} - d_{it}\rho_{it},
\]

where the expectation is taken conditional on the state vector \( z_{it} \), \( \beta \in (0,1) \) is the subjective time-discount factor, \( y \) is the household’s income, \( c \) is consumption, and \( w \) is total liquid wealth (stocks, bonds, cash accounts, and similar liquid assets). \( R^{p} \) is the real return on the risky assets held by the household and \( R^{f} \) is the return on the riskless assets, the combination of which gives the return on the portfolio \( R^{p}_{it} = \alpha_{t-1}R^{p}_{it} + (1 - \alpha_{t-1})R^{f}_{it} \). If a household decides to invest in stocks, it gives up a per-period participation cost \( \rho \). Problem (3.1) is maximized by a Markov decision rule.

Let \( V_{0}(z_{it}) \) and \( V_{1}(z_{it}) \) denote the conditional value functions associated with the decision to invest in stocks and not invest, respectively. If the household takes action \( j \) at time \( t \), the state vector \( z_{it} \) transitions into \( z_{it+1}^{(j)} \) with the probability \( F_{j}(z_{it+1}^{(j)}|z_{it}) \). The value function \( V_{j}(z_{it}) \) conditional on the choice \( j \) can be written as

\[
V_{j}(z_{it}) + \varepsilon_{jit} = \mathbb{E}_{t}\{ \sum_{\tau=t}^{T} \beta^{\tau-t}[u(z_{i\tau}, d_{i\tau}) + (1 - d_{i\tau})\varepsilon_{0i\tau} + d_{i\tau}\varepsilon_{1i\tau}]|z_{it}, d_{it} = j \}
\]

\[
= u_{j}(z_{it}) + \beta \mathbb{E}_{t}\{ p_{1}(z_{it+1}^{(j)})V_{1}(z_{it+1}^{(j)}) + p_{0}(z_{it+1}^{(j)})V_{0}(z_{it+1}^{(j)})|z_{it}, d_{it} = j \}
\]

where the second line uses the Bellman principle. The optimal participation decision rule involves comparing value functions associated with each participation choice. The optimal decision \( d_{it}^{*} \) of household \( i \) in
period $t$ can be expressed as follows:

$$d_{it}^o \equiv \begin{cases} 
1 & \text{if } V_1(z_{it}) + \epsilon_{1it} \geq V_0(z_{it}) + \epsilon_{0it} \\
0 & \text{otherwise.} 
\end{cases} \quad (3.2)$$

From equation (3.2), the conditional probability of participation in stock markets can be written as

$$p_{1it} = p_1(z_{it}) = \mathbb{E}[d_{it}^o = 1 | z_{it}] \quad \text{and} \quad p_{0it} = p_0(z_{it}) = \mathbb{E}[d_{it}^o = 0 | z_{it}].$$

Hotz and Miller (1993) prove the existence of a mapping, by which the difference in conditional value functions

$$V_1(z_{it}) - V_0(z_{it})$$

can be represented as a function of conditional probabilities. Under the assumption that $\epsilon_{jit}$ are identically and independently distributed over $(i, t)$ as Type 1 extreme value random variables, the optimality condition (3.2) can be transformed into the following:

$$\ln \frac{p_{1it}}{p_{0it}} = V_1(z_{it}) - V_0(z_{it})$$

$$= u_1(z_{it}) - u_0(z_{it})$$

$$+ \beta \mathbb{E}_t \left\{ u_0(z_{it+1}^{(1)}) - u_0(z_{it+1}^{(0)}) - \ln(p_0(z_{it+1}^{(1)})) - \ln(p_0(z_{it+1}^{(0)})) \right\} + \beta(V_{t+2}(z_{it+2}^{(1)}) - V_{t+2}(z_{it+2}^{(0)})) | z_{it}$$

The second and third lines are written using forward substitution of the conditional value function. It uses another result of Hotz and Miller (1993) that the conditional value function can be expressed in the form of the current and future utility payoffs, functions of conditional choice probabilities, and the more distant future value function $V_{t+2}$. The left-hand side of (3.3) can be estimated directly with the data. The details of the further elaboration of the right-hand side of (3.3) follow.

4 Empirical Implementation

4.1 State Variables

I limit the analysis to only a small number of participation strategies that proves sufficient for estimation of the parameters of interest. Let $x_{it}$ be an $L$-dimensional vector of exogenous covariates for household $i$ at time $t$. Income is treated as exogenous as well as forecastable and is a part of the vector of exogenous covariates. Define $\tilde{w}_{it-1} = R_{it}^o w_{it-1}$ as the period $t$ value of household $i$’s accumulated wealth. Then the observed state vector for household $i$ at time $t$ is given by the $(L + 2)$-dimensional vector $(\alpha_{it-1}, \tilde{w}_{it-1}, x_{it})$. The information set $z_{it}$ is composed of the observed state variables, and the (unobserved) individual heterogeneity statistic $v_i$
and aggregate shock \( \omega_t \), so that \( z_{it} = (\alpha_{it-1}, \tilde{w}_{it-1}, x_{it}^+, v_i, \omega_t) \). Then I define a set of histories as \((2 \times 3 + L)\)-dimensional vectors \( a_{1kit} \) and \( a_{0it} \) with the elements \((\alpha_{it-1}, \alpha_{it}, \alpha_{it+1}, \tilde{w}_{it-1}, \tilde{w}_{it}, \tilde{w}_{it+1}, x_{it}^+)\) as:

\[
a_{0it} = (\alpha_{it-1}, 0, 0, \tilde{w}_{it-1}, \tilde{w}_{it}, \tilde{w}_{it+1}, x_{it}^+),
\]

\[
a_{1kit} = (\alpha_{it-1}, \alpha_{it}, 0, \tilde{w}_{it-1}, \tilde{w}_{it}^k, \tilde{w}_{it+1}, x_{it}^+) \quad (4.1)
\]

\[
a_{1kit} = (\alpha_{it-1}, \alpha_{it}, 0, \tilde{w}_{it-1}, \tilde{w}_{it}^k, \tilde{w}_{it+1}, x_{it}^+) \quad (4.2)
\]

where \( \alpha_{it}^* \) is the optimal fraction of wealth a household chooses to invest in stocks, conditional on participating in stock markets in period \( t \). Under strategy \((4.1)\), the household chooses not to participate in the stock markets at date \( t \) (so that the corresponding state vector becomes \( z_{it+1}^{(0)} = (0, \tilde{w}_{it}, x_{it+1}^+, v_i, \omega_{t+1}) \)), and then does not participate in period \( t + 1 \). Under strategy \((4.2)\), the household chooses to participate in the stock markets at date \( t \) (with the corresponding state vector \( z_{it+1}^{(k)} = (\alpha_{it}, \tilde{w}_{it}^k, x_{it+1}^+, v_i, \omega_{t+1}) \)), but does not participate in period \( t + 1 \). If the choice of wealth held in riskless assets in period \( t + 1 \) is the same under both strategies, then the state vectors for both strategies in period \( t + 2 \), given by \( z_{it+2} = (0, \tilde{w}_{it+1}, x_{it+2}^+, v_i, \omega_{t+2}) \), do not differ neither in wealth nor in portfolio allocation. The terms in the conditional value function become inconsequential beyond time period \( t + 1 \) from the point of view of optimization.

This assumption is restrictive from the point of view of intertemporal optimization as it implies that the gains from stock market participation today are being fully consumed tomorrow. However, the assumption facilitates further inference on the magnitudes of the participation costs. Moreover, it has been implicitly maintained in the recent related empirical studies of Attanasio and Paiella (2011) and Paiella (2007), therefore, it does not diminish the incremental contribution of this study to the existing empirical papers. The important implication of this assumption is that the model can satisfy the one-period finite dependence property under the choice of participation strategy that lead to identical distribution of the state vector \( z_{it+2} \) at time \( t + 2 \). Then the future value functions beyond time \( t + 1 \) for two strategies will cancel out while taking differences in equation \((3.3)\), thus inducing finite dependence\(^9\)

Once a household decides to invest in risky assets, household wealth at time \( t + 1 \) depends on the realization of the rates of return on stock holding that is unknown to a household when it makes participation choices at time \( t \). To integrate over the uncertain returns on risky portfolio allocations, I discretize all possible realizations of the returns on stock holdings to allow for up to \( K = 3 \) states conditional on investing.

\(^9\)Hotz and Miller (1993) and Arcidiacono and Miller (2011) show that in many cases under selected participation strategies the conditional value function exhibits finite dependence, such that only a small number of future time periods matter beyond the current period. Because of the finite time dependence property, the current decisions affect only a limited number of future time periods (see Arcidiacono and Miller, 2011, for more details on finite time dependence). More distant value functions become inconsequential for estimation purposes.
generally described as high, moderate, and low realizations \((H, M, L)\). Transition probabilities are set to be independent of individual investor characteristics, so that \(F(z_{it+1}^{(k)}|z_{it}) = f_{t+1}^k\).

After taking the difference in conditional value functions associated with strategies (4.1) and (4.2), the equation (3.3) can be represented as follows:

\[
\ln \frac{p_{1it}}{p_{0it}} = u_1(z_{it}) - u_0(z_{it}) + \beta \left[ \sum_{k=1}^{K} (u_0(z_{it+1}^{(k)}) - u_0(z_{it+1}^{(0)})) - \ln(p_{0kit+1}) + \ln(p_{00it+1}) \right] f_{t+1}^k
\]

(4.3)

where \(p_{0kit+1} = p_0(z_{it+1}^{(k)})\) is the probability of not participating in stock markets in period \(t + 1\) conditional on participation in the previous time period and receiving one of the possible \(K\) realizations of return on the investment in stocks, and \(p_{00it+1} = p_0(z_{it+1}^{(0)})\) is the probability of not participating in the stock markets in period \(t + 1\) conditional on non-participation in period \(t\).

Estimation of equation (4.3) follows a three-step strategy. Step one and two estimate nuisance parameters that include estimates of fixed effects (estimates of unobserved individual heterogeneity), estimates of aggregate shocks, and estimates of conditional choice probabilities. Estimating these parameters precedes the estimation of the parameters of the utility and the participation cost. In particular, fixed effects and aggregate shocks are included in the conditioning set of the value function and conditional choice probabilities. The estimates of conditional choice probabilities are then incorporated into equation (4.3), which estimates parameters in the utility and the participation cost.

In this section, I specify the functional form of the participation costs and households’ preferences over consumption and investment choices, and then derive the equation to be estimated. Next, I discuss estimation of the nuisance parameters.

### 4.2 Participation Cost

I parameterize the participation cost \(\rho\) as a function of a household’s labor income as a monetary measure of the opportunity cost. I also allow the participation cost to be a function of age, education, and past participation choices as proxies for the experience and the ability to absorb and process financial knowledge and information. Specifically, I write participation cost as a linear combination of household’s observed characteristics:

\[
\rho_t = \delta y^t l_t x_t,
\]

where \(x_t = (1, edu_t, age_t, age_{t-1}^2, d_{t-1})\) and \(y^t\) is labor income.
4.3 Preferences

Households derive utility from the consumption good denoted by $u^c_i = u(z_i)$. I specify the utility of consumption in a quadratic form as

$$u^c_i = c_t - \xi_t c_t^2,$$

where $\xi_t$ can be a constant or a function of a household’s demographic characteristics. Quadratic utility of consumption is one of the leading models in the financial industry because of its desirable mean-variance portfolio selection properties. In the current paper, its use is appealing because it renders the estimation problem as linear. A potential drawback is that the quadratic utility does not rule out the negative marginal utility of consumption. Therefore, once the estimation of the utility parameters is achieved, ruling out parameter values that allow marginal utility of consumption to be negative for all possible values of consumption data is important.

Additionally, I allow for a non-pecuniary utility cost of adjusting the share of wealth invested in risky assets. This cost is motivated by recent evidence on the reluctance of households to adjust financial portfolio frequently. For example, Brunnermeier and Nagel (2008) argue that inertia can drive households portfolio allocation. The utility cost of adjusting portfolio composition is denoted as $u^d_i = u(d_{it}, \alpha_{it}, \alpha_{it-1})$ and specified as:

$$u^d_i = \gamma_0 d_{it} + \gamma_1 \alpha_{it} \alpha_{it-1},$$

where $\gamma_0$ is the choice specific utility shifter and $\gamma_1$ is the parameter of adjusting portfolio share invested in stocks between time $t-1$ and $t$.

Then the utility payoff for household $i$ is defined as $u_j(z_{it}) = u^c_i + u^d_i$ for $j = 0, 1$ and depends on consumption, current and past investment choices, as well as other characteristics in state vector.

4.4 The Moment Conditions

The key feature of the estimation of the participation cost is the intertemporal utility cost and benefit analysis for different participation strategies, where participation cost is connected to the amount of wealth households consider investing in stocks. On the one hand, participation cost reduces wealth today while potentially increasing household wealth in the next period, depending on the amount invested in stocks. On the other hand, households may choose not to invest in stocks today, and to keep all wealth (including the not-forgone participation cost) in risk-free assets. In the next period, however, the household can only
receive a risk-free return on savings, including the not-foregone participation cost.

For illustration, consider two identical households \( n \) and \( i \) who act under two distinct participation strategies (4.1) and (4.2), respectively. Recall that one of the strategies is to participate in period \( t \) and quit stock markets in period \( t + 1 \). The other strategy is to not participate in either period \( t \) or \( t + 1 \). Consider household \( i \), whose observed choice is to participate in period \( t \). I assume that the observed wealth is net participation cost: \( w_{it} | (d_{it} = 1) = R^p_{it} w_{it-1} + y_{it} - c_{it} - \rho_{it} \). Period \( t \) wealth of household \( n \) will be greater than the wealth for household \( i \) by the amount of the participation cost: \( (w_{nt} + \rho_{nt}) | (d_{nt} = 0) = R^p_{nt} w_{nt-1} + y_{nt} - c_{nt} \).

The identical time \( t \) state vector implies time \( t \) consumption for both households is the same; therefore, utility of consumption for these households is identical in period \( t \). Under these two strategies, however, period \( t + 1 \) consumption may not be the same, because different wealth in period \( t \) affects it. Under the first strategy, \( c_{it+1} = (\alpha_{it}(R^s_{it+1} - R^f_{it+1}) + R^f_{it+1})w_{it} + y_{it+1} - w_{it+1} \), whereas under the second strategy, \( c_{nt+1} = R^f_{nt+1}(w_{nt} + \rho_{nt}) + y_{nt+1} - w_{nt+1} \). We are unlikely to see two households with identical continuous state vectors in the data. Therefore, thinking about the same household that makes different participation choices in period \( t \) after being affected by different shocks \( \varepsilon \) can be intuitive. Looking at the transition of wealth from period \( t \) to period \( t + 1 \) under these participation strategies can illustrate their differences:

\[
\begin{align*}
\text{Strategy} & \quad (d_t = 1, d_{t+1} = 0, \ldots) & \quad \text{Strategy} & \quad (d_t = 0, d_{t+1} = 0, \ldots) \\
\quad w_t = R^p_{it} w_{it-1} + y_t - c_t - \rho_{it} & & \quad w_t + \rho_t = R^p_{it} w_{it-1} + y_t - c_t \\
\quad w_{t+1} = (R^s_{it+1} - R^f_{it+1})\alpha_t w_t + R^f_{it+1} w_t + y_{t+1} - c_{t+1} & & \quad w_{t+1} = R^f_{it+1}\rho_t + R^f_{it+1} w_t + y_{t+1} - c_{t+1}
\end{align*}
\]

It follows that under two distinct strategies, a household chooses between giving up the amount \( \rho_t \) in period \( t \) and receiving an expected increase in wealth in the amount of \( (R^s_{it+1} - R^f_{it+1})\alpha_t w_t \) next period versus not participating in period \( t \) and receiving \( R^f_{it+1} \rho_t \) next period. Finite dependence is induced by assuming that once households \( i \) and \( n \), characterized by equivalent starting conditions at time \( t \), choose not to invest in stock markets in period \( t + 1 \), their choice of wealth should not differ by much. In period \( t + 2 \), the state vector under both strategies becomes identical; therefore, the consumption and other quantities that determine value function in period \( t + 2 \) become identical as well.

After taking differences in contemporaneous utilities, the utilities from consumption at time \( t \) will cancel so that only the utility cost of adjusting one’s portfolio will remain: \( u_1(z_{it}) - u_0(z_{it}) = \gamma_0 d_{it} + \gamma_1\alpha_{it}\alpha_{it-1} \).

Under both alternatives, time \( t + 1 \) utility cost of adjusting one’s portfolio equals zero, because no investment in stocks occurs in period \( t + 1 \). The difference in utility payoffs from consumption at time \( t + 1 \)
is more involved. Denote consumption under participation strategy as \( c_{t+1}^1 \), and \( c_{t+1}^0 \) as consumption for non-participation strategy. Then the difference between utility payoffs is:

\[
\begin{align*}
    c_{t+1}^1 - c_{t+1}^0 &= -\xi_t \left[ (c_{t+1}^1)^2 - (c_{t+1}^0)^2 \right] \\
    (c_{t+1}^1 - c_{t+1}^0) \left[ 1 - \xi_t (c_{t+1}^1 + c_{t+1}^0) \right] &= \\
    (\alpha_t \hat{w}_{t+1} - R_{t+1}^f \rho_t) \left[ 1 - \xi_t (2c_{t+1}^1 + \alpha_t \hat{w}_{t+1} + R_{t+1}^f \rho_t) \right].
\end{align*}
\]

Substituting the differences in the utility payoffs into the optimality condition (4.3), I obtain the equation to be estimated:

\[
\ln \left( \frac{p_{1it}}{p_{0it}} \right) = \gamma_0 d_{it} + \gamma_1 \alpha_t \alpha_{it-1} + \beta \sum_k \left[ (\alpha_t \hat{w}_{it-k}^k - R_{it-k}^f \rho_t) \left[ 1 - \xi_t (2c_{it-k}^1 + \alpha_t \hat{w}_{it-k+1}^k + R_{it-k+1}^f \rho_t) \right] - \ln \frac{p_{0it-k+1}}{p_{00it-1}} \right] \gamma_{it-k},
\]

where \( \hat{w}_{it-k}^k = w_{it}(R_{it-k}^f - R_{it-k+1}^f) \) is a function of the data.

Let \( \Theta \) denote all unknown parameters in the model to be estimated. These parameters include the utility parameters \( \gamma_0, \gamma_1, \) and \( \xi \) and the parameters of the participation cost \( \delta \). I fix the value of time discount factor \( \beta \) and estimate the remaining parameters conditioning on \( \beta \). I rearrange and combine terms represented by observables to get the equation:

\[
m_{it}(\Theta) = Y_{it} - X_{it} \Theta, \quad (4.4)
\]

where

\[
Y_{it} = \ln \left( \frac{p_{1it}}{p_{0it}} \right) - \beta \sum_k \left[ \alpha_t \hat{w}_{it-k+1}^k - \ln \frac{p_{0it-k+1}}{p_{00it-1}} \right] \gamma_{it-k},
\]

\[
X_{it} = \left( 1, \quad \alpha_t \alpha_{it-1}, \quad \beta \rho_{it}^j y_{it}^j, \quad \beta \sum_k \alpha_t \hat{w}_{it-k+1}^k (2c_{it-k}^1 + \alpha_t \hat{w}_{it-k+1}^k) \gamma_{it-k}, \quad 2\beta \rho_{it}^j (\sum_{j=1}^{L-1} s_{it-j}^j, \quad \beta \rho_{it}^j \sum_{j=1}^{L-1} s_{it-j}^j, \quad 2\beta \rho_{it}^j \sum_{j=1}^{L-1} s_{it-j}^j, \quad \beta \rho_{it}^j \sum_{j=1}^{L-1} s_{it-j}^j). \right)
\]

where \( L \) is the dimension of \( x_{it} \). \( \Theta = (\Theta_1, \Theta_2) \), where \( \Theta_1 = (\gamma_0, \gamma_1, \delta, \xi) \) has dimension \( 2 + L + H \), and \( \Theta_2 = (\xi \delta, \xi \delta^2, ..., \xi \delta_{L-1} \delta_L) \) has dimension \( 2HL + HL(L - 1)/2 \), where \( \delta \) is a vector of parameters \( (\delta_1, ..., \delta_L) \) with the same dimension as \( x_{it} \) and \( H \) is the dimension of \( \xi \). My primary interest is only in \( \Theta_1 \). In the estimation, I do not impose any restrictions on \( \Theta_2 \) and estimate the unrestricted equation (4.4) that is linear in parameters.

If I were to observe or estimate the conditional choice probabilities and the transition probabilities, I
could estimate the unknown structural parameters of interest from a conditional moment condition

\[ E[m(X_{it}, Y_{it}, \Theta_o)|z_{it}] = 0, \] (4.5)

where subscript \( o \) denotes the true value of the parameters. The minimum distance estimator is a natural estimator choice in this and similar frameworks.

### 4.5 Individual Effects and Aggregate Shocks

Microeconometric literature has proposed different methods for incorporating unobserved heterogeneity in dynamic discrete choice models. Altug and Miller (1998) estimate fixed effects from an auxiliary regression related to the main model. More recent studies by Aguirregabiria and Mira (2007), Kasahara and Shimotsu (2009), and Arcidiacono and Miller (2011) propose alternative ways through finite mixture distributions. The approach of Kasahara and Shimotsu is restrictive for my estimation. For identification of unobserved heterogeneity, their approach requires the time dimension of a panel that is beyond the one available in the data at hand. The identification requirements in Arcidiacono and Miller are similar to those in the study of Kasahara and Shimotsu. I adopt the approach of Altug and Miller and use the earnings equation to estimate individual unobserved effects.

The same earnings equation allows me to estimate time effects as well. In my framework, current wealth accumulation is affected by aggregate shocks through current wages and rates of return. Aggregate shocks on rates of return are captured by the risk-free rate, \( R^f_t \). Aggregate shocks on wages are not directly observed, but can be estimated.

I assume both time effects and individual effects impact a household’s earnings are affected by both time effects and individual effects; hence I can estimate these effects by modeling the household’s earnings process. I consider a dynamic earnings equation of the form:

\[ \hat{y}_{it} = \phi\hat{y}_{i,t-1} + x_{it}'\kappa + \omega_t + v_i + e_{it}, \] (4.6)

where \( \hat{y}_{it} \) is the log-transformation of a household’s labor income and \( x_{it}' \) is an \( L \)-dimensional vector of family characteristics for household \( i \) at time \( t \), \( \omega_t \) is the unobserved time-specific effect, and \( v_i \) is the unobserved individual-specific effect. I follow Arellano and Honore (2001) in assuming the predeterminedness condition of \( \hat{y}_{i,t-1} \) and (possibly) \( x_{it} \), precisely \( E(e_{it} | \hat{y}_{i,t-1}^{t-1}, x_{it}^t) = 0 \) and \( E(v_i) = 0 \). Subject to a rank condition, \( (1 + L + \)
parameters of the model (4.6) are identified with $T \geq 3$. With $T = 3 \phi, \kappa', \omega_2$, and $\omega_3$ are just identified from the $3 + L$ orthogonality conditions:

$$E \left[ \begin{pmatrix} 1 \\ x'_{it} \\ \hat{y}_{it-2} \\ \Delta \hat{y}_{it} - \phi \Delta \hat{y}_{it-1} - \Delta x'_{it} \kappa - \Delta \omega_t \end{pmatrix} \right] = 0 \quad (4.7)$$

$$E \left[ \Delta \hat{y}_{it} - \phi \Delta \hat{y}_{it-1} - x'_{it} \kappa - \omega_t \right] = 0,$$

Estimation of the fixed effects, however, requires a larger $T$. The extended data set covers $T = 14$ time periods of observations. The vector of family characteristics $x'_it$ contains three variables - family size, education multiplied by age and age squared. It makes 17 parameters to estimate. The overidentified system of equations (4.6) contains four moment conditions for each of $T - 2$ time periods, producing up to 48 orthogonality conditions for each household. The four moment conditions include the unconditional moment condition in levels and the moment conditions in first differences with a constant, income at time $t - 2$, and age squared at time $t$ as instruments. I estimate the system of equations (4.7) with the GMM procedure.

Table 2 reports the estimates of the earning equation (4.6). The estimates on family size, education, and age are consistent with the general expectations. Household income increases with family size. It responds positively to the level of education and decreases with the household growing older. The $AR(1)$ coefficient on past earnings is estimated at 0.85 and is statistically significant. Most estimates on $\Delta \omega_t$ are also statistically significant. With the parameter estimates of model (4.6), individual fixed effects are trivially estimated.

4.6 Conditional Choice Probabilities

Following the methodology developed in Hotz and Miller (1993) and Altug and Miller (1998), I compute the conditional choice probabilities nonparametrically using a kernel estimator. The conditional choice probability $p_j(z_{iklt})$, $j,l = 0,1$ and $k = 0,...,K$, is denoted as a probability of choosing the participation alternative $j$ by household $i$ in period $t$ given that the alternative $l$ was chosen in period $t - 1$ and conditioned on realization of the state $z_{iklt}$. Table 3 reports the summary of nonparametric estimates of conditional choice probabilities. The table also reports the number of observations used to compute the probability associated with each relevant history $z_{iklt}$ as well as sample standard deviations. The estimated conditional choice probabilities are limited to those relevant to the histories employed to derive equation (4.3).

Line one of Table 3 reports the participation rate $p_{i1t}$. Line 2 reports the probability of not participating
Table 2: Earnings equation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std.Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.856</td>
<td>0.038</td>
</tr>
<tr>
<td>$\kappa's$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>family size</td>
<td>0.202</td>
<td>0.037</td>
</tr>
<tr>
<td>$edu^\times age$</td>
<td>0.577</td>
<td>0.063</td>
</tr>
<tr>
<td>$age^2$</td>
<td>-0.863</td>
<td>0.088</td>
</tr>
<tr>
<td>$\omega_{1983}$</td>
<td>-0.720</td>
<td>0.357</td>
</tr>
<tr>
<td>$\Delta\omega_{1985}$</td>
<td>0.078</td>
<td>0.015</td>
</tr>
<tr>
<td>$\Delta\omega_{1987}$</td>
<td>-0.063</td>
<td>0.014</td>
</tr>
<tr>
<td>$\Delta\omega_{1989}$</td>
<td>-0.025</td>
<td>0.013</td>
</tr>
<tr>
<td>$\Delta\omega_{1991}$</td>
<td>-0.088</td>
<td>0.013</td>
</tr>
<tr>
<td>$\Delta\omega_{1993}$</td>
<td>0.012</td>
<td>0.013</td>
</tr>
<tr>
<td>$\Delta\omega_{1995}$</td>
<td>-0.108</td>
<td>0.016</td>
</tr>
<tr>
<td>$\Delta\omega_{1997}$</td>
<td>0.031</td>
<td>0.017</td>
</tr>
<tr>
<td>$\Delta\omega_{1999}$</td>
<td>0.084</td>
<td>0.018</td>
</tr>
<tr>
<td>$\Delta\omega_{2001}$</td>
<td>-0.025</td>
<td>0.015</td>
</tr>
<tr>
<td>$\Delta\omega_{2003}$</td>
<td>-0.071</td>
<td>0.015</td>
</tr>
<tr>
<td>$\Delta\omega_{2005}$</td>
<td>0.057</td>
<td>0.017</td>
</tr>
<tr>
<td>$\Delta\omega_{2007}$</td>
<td>-0.008</td>
<td>0.019</td>
</tr>
</tbody>
</table>

$J$ test: 10.5
$p$ value: 0.99

1 Number of time periods $T = 14$, number of households in the sample 7,744. Instruments include a constant, age of head of household squared at time $t$, and labor income at time $t - 2$.

Table 3: Sample averages of nonparametric estimates of conditional choice probabilities

<table>
<thead>
<tr>
<th>Num.obs.</th>
<th>Sample mean</th>
<th>Sample std.dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $p_{1it}$</td>
<td>10,639</td>
<td>0.256</td>
</tr>
<tr>
<td>2. $p_0(z_{0it})$</td>
<td>9,412</td>
<td>0.889</td>
</tr>
<tr>
<td>3. $p_0(z_{Hit})$</td>
<td>10,590</td>
<td>0.030</td>
</tr>
<tr>
<td>4. $p_0(z_{Mhit})$</td>
<td>10,588</td>
<td>0.016</td>
</tr>
<tr>
<td>5. $p_0(z_{Lhit})$</td>
<td>10,608</td>
<td>0.138</td>
</tr>
</tbody>
</table>
in the stock markets in period \( t \) conditional on not participating in period \( t - 1 \). Lines three, four, and five report the probability of not participating in stock markets in period \( t \) conditional on participation in the previous time period and receiving high, moderate, or low return on the investment in stocks. The average estimated participation rate is 0.25. The results also suggest the probability of participation decreases if a previous period has an incidence of nonparticipation to 0.11.

5 Empirical Findings

Table 4 contains the estimation results for the utility parameters and parameters of the participation cost. Columns (1)-(3) report estimates for the model with the simple quadratic utility of consumption, columns (4)-(6) report estimates for the model with age-augmented quadratic utility of consumption. The table reports the total number of parameters estimated by each model specification, but only includes \( \gamma_0, \gamma_1, \delta', \) and \( \xi' \)\(^{10}\)

**Participation cost**

Columns (1) and (4) of Table 4 show the participation cost is estimated on average to be about 4% to 6% of labor income. The participation cost measured as a share of income is not constant over the life cycle and admits the curvature of the labor income profile. Results in columns (2)-(3) and (5)-(6) provide evidence that, even holding labor income fixed, age, education, and previous participation affect participation costs. Table 4 shows the participation cost is decreasing in education. It is roughly decreasing with age. When I introduce the curvature in age in the participation cost, it becomes increasing and concave with age, with the turning point around ages 35-40. Finally, the participation cost is smaller for households that participated previously in stock markets.

The estimation results for the participation cost seem to be intuitive and agree with previous findings. First, Gomes and Michaelides (2003) rationalize limited participation by the existence of the participation cost. Second, empirical findings of Bertaut and Starr-McCluer (2002), Guiso and Japelli (2002), and Banks and Tanner (2002) show individual characteristics such as age and education can explain participation in stock markets. Finally, past participation also affects future participation. Alessie, Hochguertel, and van Soest, (2002, 2004), and Munoz (2006) empirically confirm this dynamic dependence. The estimation results in the current study relate to these findings and show the participation cost is significantly different from zero and varies with household characteristics.

\(^{10}\)The full set of parameter estimates is available from the author.
Table 4: Participation cost and utility parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility parameters $(c_t - \xi_t^2) + \gamma_0 d_t + \gamma_1 \alpha_t \alpha_{t-1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi_t = \xi_1$</td>
<td>$\xi_t = \xi_1 + \xi_2 \text{age}_t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>24.414 (3.829)</td>
<td>34.174 (4.653)</td>
<td>33.183 (5.403)</td>
<td>20.588 (3.852)</td>
<td>30.590 (4.896)</td>
<td>25.337 (7.122)</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>11.535 (7.739)</td>
<td>−5.906 (8.998)</td>
<td>−5.344 (9.680)</td>
<td>11.112 (7.898)</td>
<td>−0.501 (7.082)</td>
<td>2.110 (8.336)</td>
</tr>
<tr>
<td>$\xi_1$</td>
<td>0.00051 (0.00015)</td>
<td>0.00046 (0.00015)</td>
<td>0.00046 (0.00015)</td>
<td>−0.00028 (0.00071)</td>
<td>−0.00036 (0.00084)</td>
<td>−0.00009 (0.00083)</td>
</tr>
<tr>
<td>$\xi_2$</td>
<td>0.000016 (0.000013)</td>
<td>0.000019 (0.000016)</td>
<td>0.000015 (0.000016)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participation cost $\rho_t = y_t (\delta_1 + \delta_2 \text{edu} + \delta_3 \text{age}<em>t + \delta_4 \text{age}^2_t + \delta_5 d</em>{t-1})$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.061 (0.019)</td>
<td>0.377 (0.093)</td>
<td>0.043 (0.221)</td>
<td>0.042 (0.019)</td>
<td>0.352 (0.108)</td>
<td>0.0344 (0.3124)</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>−0.0163 (0.0049)</td>
<td>−0.0180 (0.0050)</td>
<td></td>
<td>−0.0158 (0.0058)</td>
<td>−0.0157 (0.0072)</td>
<td></td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>−0.0007 (0.0009)</td>
<td>0.0170 (0.0096)</td>
<td></td>
<td>−0.0005 (0.0011)</td>
<td>0.0146 (0.0130)</td>
<td></td>
</tr>
<tr>
<td>$\delta_4$</td>
<td></td>
<td></td>
<td></td>
<td>−0.00021 (0.00011)</td>
<td></td>
<td>−0.00019 (0.00015)</td>
</tr>
<tr>
<td>$\delta_5$</td>
<td>−0.0246 (0.0137)</td>
<td>−0.0284 (0.0139)</td>
<td></td>
<td>−0.0322 (0.0154)</td>
<td>−0.0363 (0.0194)</td>
<td></td>
</tr>
<tr>
<td>Number of parameters</td>
<td>6</td>
<td>20</td>
<td>26</td>
<td>9</td>
<td>34</td>
<td>46</td>
</tr>
<tr>
<td>$J$-test</td>
<td>98</td>
<td>112</td>
<td>125</td>
<td>107</td>
<td>134</td>
<td>138</td>
</tr>
<tr>
<td>$p$-value</td>
<td>5e-7</td>
<td>4e-4</td>
<td>7e-4</td>
<td>5e-7</td>
<td>0.0051</td>
<td>0.087</td>
</tr>
</tbody>
</table>

Standard errors in parenthesis. Standard errors are estimated using a bootstrap method with 200 replications.
Figure 1: **Age profile of the participation cost as a share of labor income.** Education and past participation are fixed at the sample average level.

Figure 1 illustrates how participation cost (as a share of labor income) changes over age, with other demographic characteristics fixed at their sample average levels. Panel A presents the age profile of the participation cost computed using parameter estimates from Table 4 columns (1)-(3), while panel B presents the same for columns (4)-(6). Participation cost can be as great as 10% of labor income, but also can decrease to negligible values. On average, participation cost in columns (1)-(3) is greater than in columns (4)-(6) by about 2 percentage points of labor income. The effect of age on both the participation cost and the utility can explain these differences. The specification in columns (4)-(6) is different from the one in columns (1)-(3) only through the age-dependent parameter $\xi_t$ in the utility of consumption. Once the age is controlled for in the curvature of the utility of consumption, the effect of age on participation cost decreases.

Overall, the estimation results provide additional support for the empirical findings of Paiella (2007) and Attanasio and Paiella (2011). The estimated participation cost is on average much greater than the lower bound estimates in these studies. Also, the estimated per period participation costs support the existence of fixed entry costs used in the theoretical papers of Gomes and Michaelides (2003), Haliassos and Michaelides (2003) and related studies. Column (6) suggests per-period participation costs are substantially smaller (by 3.6%) for investors with previous stock market participation experience. This finding implies first time-participants face larger participation costs that potentially include a fixed entry cost.

On the one hand, if a consumer faces only fixed entry costs while making a decision to invest in stocks,
then once he decides in favor of investing, quitting the stock market and returning later is not costly for him. In this case, he has already incurred the cost and the information he obtained when entering the stock market the first time, never depreciates. On the other hand, participation cost can measure the ongoing effort to gather the information related to investment in stocks, and the knowledge an investor gained in previous periods may still be relevant in the current period. However, the longer ago an investor obtained the knowledge, the less applicable and relevant it may be to the current state of the stock market. The per-period participation cost can, therefore, capture the effect of the diminishing relevance of previous knowledge and experience.

Because of the limitations along the time dimension of the panel data, I allow for one lag of the participation dummy to affect the current participation cost. In general, one can test the relevant importance of the entry cost and per-period cost with the flexible form of the participation cost if we allow more lags in the participation. This allowance is possible at the expense of the reduction of the time dimension in the panel used in estimation. I defer the extension of the model along this line of research for future work.

**Utility parameters**

Utility parameters include the curvature parameters in the utility of consumption, the utility shifter for stock market participants, and parameters of the adjustment of portfolio allocation. The estimates of $\xi_t$ are significant in all model specifications and indicate the utility of consumption is concave. The average value of $\xi_t = \xi_1 + \xi_2 \text{age}_t$ is similar to the estimates of $\xi_t$ in columns (1)-(3). Even with the negative $\xi_1$ in columns (5) and (6), $\xi_t$ never takes negative values. Additionally, the estimates imply a positive marginal utility of consumption for all consumption values in the data.

All model specifications in Table 4 suggest the existence of a positive utility shifter for the households that participate in stock markets. The parameter of the portfolio adjustment cost is not significantly different from zero. The positive estimate of $\gamma_1$ would indicate a complimentarity effect of past portfolio allocation on current portfolio allocation, and would also provide evidence on the effect of inertia, found by Brunnermeier and Nagel (2008). However, the current estimation outcome does not allow a revelation of such effects in a precise manner.

The $J$ tests associated with the model specifications reported in Table 4 indicate the rejection of over-identifying restrictions implied by the choice of instruments for specifications (1) through (5). However, the model specification (6) cannot be rejected at the 5% significance level. I use the parameter estimates from this specification to evaluate how the model fits to the data.
6 Simulation Results from the Structural Model

6.1 Model Fit

The simulation exercise in this section is aimed to show how the estimated model fits the stock market participation pattern observed in the data. I solve the model via backward induction using the estimated parameters reported in column (6) of Table 4. As the estimation uses biannual data, I also solve the model recursively on a biannual basis. To align with the estimation set-up that considers households prior to retirement (i.e. before age 65), I simulate households up to the retirement stage.

To solve for the optimal share of wealth invested in risky assets and total wealth, I first define a fine grid on state variables that include past share of wealth invested in risky stocks and past wealth with earned rate of return, as well as income. I achieve the interpolation off the grid with the second order polynomial regression of the value function at each point of the grid on the corresponding state space vector. This regression provides an accurate approximation to the value functions off the grid with $R^2$ in the range of 0.98-0.99.

Once I obtain the solution, I simulate the model, producing 5,000 replications. I choose no participation in stock markets as a starting condition for all simulated individuals. Conditional on the state variables, I compute the conditional choice value function for each combination of the choice variables from the grid, using the interpolation coefficients obtained at the solution as well as the draws of extreme value random variables. Then I select the maximal conditional choice value function and find the optimal choice of participation decision, portfolio choice, and wealth associated with the maximal value function. If a simulated consumer finds participating in stock markets is optimal, he receives the rate of return of stock holdings, drawn randomly as low, moderate, or high return with corresponding probabilities used in estimation as transition probabilities.

Figure 2 compares (1) the average age-specific participation rate computed from the data, (2) participation rate, predicted by the probit model, which I estimated earlier in the section on the data description, and (3) participation rate, predicted by the structural model. The participation rate generated by the probit model is reported for comparison. The probit model employs a somewhat richer set of covariates, including time dummies and cohort effects, in a simpler and flexible model structure that, however, does not allow for explicit modeling of participation costs. As shown in Figure 2, the structural model with its more complex

\[11\] A higher order polynomial regression was also tried; however, it produces less accurate predictions.

\[12\] A greater number of simulations makes little difference.
Figure 2: Participation rate: Data vs Simulation
structure and explicit accounting for participation costs successfully predicts the stock market participation rate observed in the data. The model also produces the increasing pattern of stock market participation over the life cycle.

6.2 Robustness Check

The estimation of the participation cost relies on the utility of consumption to be quadratic. Quadratic preferences, while being analytically and computationally convenient, have arguably undesirable risk aversion characteristics. In this section, I conduct a robustness check aimed to evaluate how the estimated heterogeneous participation cost fits with the alternative preferences. In particular, CRRA utility is considered as the alternative, because isoelastic preferences have been widely used in the literature on stock market participation choices. Some examples among the recent applications of CRRA preferences include the papers on portfolio choices of the wealthy consumers of Carroll (2002) and Guvenen (2007), calibration of life cycle portfolio choices under uninsurable labor income risk and borrowing constraints (Cocco, Gomes and Maenhout, 2005) and under the house price risk (Cocco, 2004), studies that reveal preference heterogeneity.
among stock market participants and non-participants (Attanasio, Banks and Tanner, 2002; Brav, Constantinides, and Geczy, 2002; Vissing-Jorgensen, 2002a) and papers that address the issue of the stock market participation costs (Alan, 2006; Paiella, 2007; Attanasio and Paiella, 2011). The parameter of the relative risk aversion (RRA) is chosen to be 1.7. This value of the RRA parameter is the result from the recent microdata estimation of Attanasio and Paiella (2011) for the sample of the US consumers, both stock market participants and non-participants. This value is well within the existing range of the empirical estimates of the RRA. Other values of the RRA have been tried in a wider range between 1 and 5 with no substantial change in the reached conclusion, but with the resulting assertion that the chosen value indeed fits the data well.

Figure 3 plots the stock market participation rate implied by CRRA preferences together with the participation rate implied by the estimated benchmark model and the average age-specific participation rate computed from the data. On average, the simulated participation rate implied by CRRA preferences matches the participation rate observed in the data. Starting around age 50 the simulated participation rates under CRRA and quadratic preferences are extremely similar, however, the two series give different stock market participation predictions at the younger age. The participation rate implied by CRRA preferences stays relatively constant over the life cycle and does not admit the increasing life cycle pattern of stock market participation observed in the PSID. The relatively stable participation rate of a similar magnitude is also achieved in Gomes and Michaelides (2005) for the case of preference heterogeneity.

Overall, the model with the estimated heterogeneous participation cost and CRRA preferences generates the average participation rate similar to the one observed in the data and previously reported in the literature.

7 Conclusion

In this paper, I attempt to rationalize the limited stock market participation at the micro-data level. I develop and estimate a dynamic structural model of individual stock market participation and portfolio choice, where decisions regarding stock market participation are influenced by participation costs. The practical significance of the participation costs is considered as being a channel through which financial education programs can affect consumers’ investment decisions. Financial education and counseling alleviates the burden on consumers’ time and the effort necessary for making financial decisions and reduces the objective cost of stock market participation. The estimation results support the existence of a significant cost of stock market participation that, on average, is about 5% of labor income. These results augment the findings on
the lower bound of the stock market participation cost currently available in the empirical literature, but also are in accordance with the fixed entry costs used in the theoretical literature.

The estimated model not only supports empirically the strand of the theoretical and empirical literature that utilizes participation costs to explain limited stock market participation, but also facilitates simulated paths for the decision to participate in stock markets that successfully line up with patterns observable in the data. While the theoretical models of stock market participation with participation costs are able to generate lower participation rates, the predicted levels of participation in stock markets are still substantially greater than those observed in the data. The striking accuracy of the estimated structural model is likely due to its richer stochastic environment and more sophisticated accounting for individual heterogeneity.

Overall, this paper is the first analysis of the stock market participation and portfolio composition that allows for heterogeneity in the participation cost and estimates the magnitude of it using micro data. The empirical evidence suggest the participation costs decrease with stock market participation experience and as consumers become older, and are less for more educated investors. In general, the results contribute to the broad strand of the literature on life-cycle consumption and savings decisions, as well as on the idiosyncratic response to consumer financial protection policy initiatives.
References


