Portfolio Choice and Asset Pricing with Investor Entry and Exit*

Yosef Bonaparte, George M. Korniotis, Alok Kumar†

May 6, 2018

Abstract

We find that about 25% of stockholders enter/exit non-retirement investment accounts biennially. To examine this participation turnover, we estimate a canonical portfolio choice model. The estimation reveals that income risk and time costs to investing, estimated at 3.7% of income, affect the turnover in stock ownership. Further, the estimates of the coefficient of relative risk aversion and the discount factor are 3.176 and 0.963, respectively. Finally, the model implies that consumption CAPMs can perform better when focusing on investors who rarely exit the market because their consumption growth is the most volatile and the most correlated with market returns.

Keywords: Consumption risk, limited stock market participation, trading costs, income risk, PSID, SCF.

JEL classification: D14, G11, G12.

---

*We are thankful for comments and suggestions from Sandro Andrade, William Bazley, Russell Cooper, Stefanos Delikouras, Sarah Khalaf, and Sheridan Titman. All remaining errors and omissions are our own.

†Bonaparte is at the University of Colorado Denver, 1475 Lawrence St., Denver, CO 80202, USA and can be reached at yosef.bonaparte@ucdenver.edu. Korniotis (gkorniotis@miami.edu), and Kumar (akumar@miami.edu) are at the Department of Finance, School of Business Administration, 514 Jenkins Building, University of Miami, Coral Gables, FL 33124, USA. Please address all correspondence to George Korniotis.

1 Introduction

Stock market participation is an important portfolio choice decision. Accordingly, the literature studies a wide array of issues related to stock ownership. They range from characterizing the demographics of stock holders (Campbell 2006) and their investment mistakes (Calvet, Campbell, and Sodini 2007) to estimating preference parameters for stockholders, such as the elasticity of intertemporal substitution (Vissing-Jørgensen 2002a).

However, with the exception of a few papers[1] the prior literature has not carefully studied the dynamic decision to enter/exit the market. Moreover, the implicit assumption in much of the empirical asset pricing literature is that after an investor enters the market, he stays in the market for the long term. This assumption might be true for retirement accounts, but little is known on whether it holds for non-retirement investment accounts. Overall, there is little evidence about the frequency of entry and exit in the stock market and the unique factors determining the overall turnover in stock market participation.

To fill this gap in the literature, we examine the dynamic decision to own stocks empirically and theoretically. To begin, we provide empirical evidence of large turnover in and out of non-retirement accounts. Then, we formally examine the dynamic participation decision within a portfolio choice model with borrowing constraints, income shocks, and transaction costs. We estimate the model and show that it captures relatively well the average participation rates and the average entry/exit rates. Finally, simulated data from the model indicate that the consumption growth risk of stockholders depends on their participation history, where stockholders with low exit rates have the highest consumption risk. Also, the correlation of their consumption growth with market returns is the highest.

We begin our analysis by presenting some new stylized facts related to the entry and

[1] For example, see Hurst, Luoh, and Stafford (1998), Vissing-Jørgensen (2002b), Calvet, Gonzalez-Eiras, and Sodini (2004), Calvet, Campbell, and Sodini (2009), and Bilias, Georgarakos, and Haliassos (2010).

exit decisions of investors. In particular, we use the biennial waves of the Panel Study for Income Dynamics (PSID) from 1999 to 2011. We focus on households who own stocks directly and indirectly. Because we want to examine active stock ownership, our stockholder classification excludes households who only own stocks via retirements accounts such as employer-administered pension funds and IRAs. We exclude investments in retirement accounts because there is little trading in these accounts (Ameriks and Zeldes 2004a).[2]

We find that on average among all the PSID households in a given wave, about 7.3% (8.7%) of them enter (exit) the stock market over the next two years. Among the current stockholders (non-stockholders) about 28.8% (23.8%) on average exit (enter) the market by the next survey wave. Furthermore, only 32.8% of households that owned stocks in 1999, reported owning stocks in all subsequent waves until 2011.

We obtain similar evidence using the Survey of Consumer Finances (SCF). In the 2007-2009 SCF panel, among the households who participated in the market in 2007, 13.8% exited by 2009. Moreover, among the households who did not participate in the market in 2007, 20% of them entered the market in 2009. Consistent with evidence on trading by retail investors (Barber and Odean 2000; Barber and Odean 2001), these statistics demonstrate strong dynamics in the stock market participation decision.

We use a portfolio choice model to examine the factors affecting the turnover in market participation. To set the stage for the model, we estimate regressions that directly test the main intuition of portfolio choice models related to market participation. As suggested by Guiso, Jappelli, and Terlizzese (1996), canonical model with stochastic income imply that investors facing income risk and severe liquidity constraints would trade the most.

We confirm this intuition by estimating cross-sectional regressions where the dependent variables are related to the total number of entries and exits from the stock market. Our

[2]See Appendix A for a description of the various stockholder classifications used in the literature.

main independent variables are the standard deviation of income growth (proxing for income risk) and wealth and income (proxing for liquidity constraints). Our control variables are age, gender, race and education. We estimate the regressions using the PSID waves from 1999 to 2011. Consistent with our intuition, we find that the coefficient estimates on income risk are positive and the coefficient estimates on wealth and income are negative.

Next, we build a portfolio choice model that explicitly allows for entry and exit from the stock market. In the model, there are ex-ante homogeneous households who receive an exogenous stochastic income payment. We allow for income risk because the household finance literature finds that it affects portfolio decisions.[3] Based on income and wealth, the households decide how much to consume and how much to save in a risky and a risk free asset. The households are also subject to short-sale and borrowing constraints.

An important assumption in the model is that households face costly entry and exit from the stock market and incur transaction costs when they trade. Following the literature, we assume that they can incur three types of transaction costs.[4] The first one is a fixed cost to trading capturing expenses such as investment account maintenance fees. The second one includes variable trading costs accounting for brokerage fees, commissions, and the bid-ask spread. The third cost is a per-period time cost to trading that we model as lost income.

We estimate the model using the simulated method of moments (SMM). The estimation includes household-level moments related to the frequency of entry/exit, portfolio adjustment rates, and wealth ratios. The estimation also includes market-level moments related to the equity premium and the reaction of aggregate consumption growth to returns. We use aggregate moments in the estimation to ensure that the aggregate implications of the

[3]See Guiso, Jappelli, and Terlizzese (1996), Angerer and Lam (2009), Betermier, Jansson, Parlour, and Walden (2012), Bonaparte, Korniotis, and Kumar (2014), and Betermier, Calvet, and Sodini (2017).

[4]See Luttmer (1999), Campbell, Cocco, Gomes, and Maenhout (2001), Paiella (2001), Vissing-Jørgensen (2002b), Calvet, Gonzalez-Eiras, and Sodini (2004), Paiella (2004), Alan (2006), Paiella (2007), Attanasio and Paiella (2011), and Bonaparte, Cooper, and Zhu (2012).

household-level model are not inconsistent with market-level stylized facts.

In the SMM estimation, we focus on estimating the deep parameters of the model like the coefficient of relative risk aversion and the time cost to investing. Similar to Bonaparte, Cooper, and Zhu (2012), we do not estimate within the SMM the variable trading costs or the household income process. Instead, we estimate the variable cost function directly using the brokerage investor data set of Barber and Odean (2000). We also estimate the household-level income process directly using data from the PSID.[5]

The estimation reveals that our model can capture relatively well the decision to enter/exit the stock market and it is not rejected by the J-test of over-identifying restrictions. More importantly, the estimates of the preference parameters are reasonable. For example, the estimated coefficient of the relative risk aversion and the discount factor are 3.176 and 0.963, respectively. These estimates are close to those in the literature (Cagetti 2003; Bonaparte, Cooper, and Zhu 2012). Our model also produces a reasonable equity premium as found in other studies with limited stock market participation (Attanasio and Paiella 2011).

After the estimation, we use the model to examine the impact of various transaction cost components on the dynamic decision to own stocks. We find that fixed-type participation costs and proportional trading costs have little effect on the entry/exit decisions. The transaction cost that is the most important is the per-period participation cost representing the time cost to trading. We estimate this cost to be 3.20% of current income. Using the average annual income in our PSID sample ($72,000), this cost is about $2,304.

Even though the time cost estimate seems high, it is consistent with the current costs of delegating investment decisions. In particular, Wermers (2000) finds that the average expense ratio of active mutual funds is around 0.93% of assets under management. We

[5]For more details, please see Appendix B for the estimation of the cost function and Appendix C for the estimation of household income process.

use this expense percentage together with the wealth of the average household invested in equity from the SCF, and find that the implied average delegation cost ranges from \$1,471 to \$2,154. These back-of-the envelope estimates are close to our average time cost of \$2,304.

Next, we examine the sources of risk that can impact the investors' dynamic decision to trade. For this analysis, we estimate cross-sectional regressions using simulated data from the model and find that households with high income risk exhibit high frequencies of entry and exit from the stock market. We also estimate time-series regressions and find that higher stock market returns decrease the aggregate frequency of exit and increase aggregate stock market participation. Also, an increase in the growth of national average income, increases (decreases) the frequency of entry (exit).

Overall, the most important determinants of the entry and exit decisions are time costs related to trading as well as stock market and income shocks. More broadly, our findings show that a simple canonical model of portfolio decisions can simultaneously fit participation decisions and the equity premium with realistic preference parameter estimates.

Our work is related to the literature on the behavior of retail investors. This literature focuses on trading of stocks and puts less emphasis on the entry/exit aspects of trading (Barber and Odean 2000; Barber and Odean 2001). To explain stock trading behavior, Grossman and Stiglitz (1980) suggest that investors equate the marginal benefit of trading to its marginal cost. Based on this principle, a large literature examines the impact of transaction costs on portfolio decisions.[6] In contrast, Odean (1998) suggests that investors are overconfident, which leads to excessive trading and lower portfolio performance.

Other related research finds that households adjust their portfolio composition in response to changes in wealth, income, and age, as well as stock market performance and

[6] Among others, see Constantinides (1976), Constantinides (1986), Dumas and Luciano (1991), Gennotte and Jung (1994), Longstaff (2001), Liu and Loewenstein (2002), and Gârleanu and Pedersen (2013).

volatility.[7] For example, Bilias, Georgarakos, and Haliassos (2010) show that young, white, healthy, college graduates with high income and high wealth trade the most. Calvet, Campbell, and Sodini (2009) find that the probability of entry (exit) is higher (lower) for households with higher (lower) wealth, income and education. Barrot, Kaniel, and Sraer (2016) find that during the 2008 crisis, French retail investors provided liquidity to the market by selling stock funds and buying individual stocks. Dorn and Weber (2017) show that households who delegate their investments have a higher probability of exiting the market during times of crisis relative to households that directly own stocks.

Calvet, Gonzalez-Eiras, and Sodini (2004) also build a theoretical model with endogenous participation and heterogeneous income risks. They use the model to study the impact of financial innovation on the stock ownership turnover. They find that the introduction of new assets reduces the income hedging costs for some households who are encouraged to enter the stock market. But, increased participation results into lower risk premia thus reducing the profitability of investments for some investors who ultimately exit the market. Complementing their finding, we also argue that mitigating income risk is a primer driver of the dynamic decision to own stocks.

Our results have important implications for asset pricing tests that use pricing factors related to stockholders. For example, Mankiw and Zeldes (1991) find that the CCAPM performs better when using the consumption growth of stockholders. Attanasio, Banks, and Tanner (2002) find that the estimate of relative risk aversion is reasonable and precisely estimated from a sample of likely stockholders. Vissing-Jørgensen (2002a) finds significant estimates of the elasticity of intertemporal substitution (EIS) when estimating Euler equations of U.S. households that either own stocks or bonds. Brav, Constantinides, and Geczy

[7] See Haliassos and Bertaut (1995), Heaton and Lucas (1996), Bertaut and Haliassos (1997), Bertaut (1998), Gollier (2001), Viceira (2001), Campbell and Viceira (2002), Haliassos and Michaelides (2003), Cocco, Gomes, and Maenhout (2005), and Gomes and Michaelides (2008).

(2002) find that the equity and value premia can be rationalized with a stochastic discount factor based on the consumption growth of market participants.

More recently, Malloy, Moskowitz, and Vissing-Jørgensen (2009) highlight the importance of the long run consumption risk of stockholders in fitting expected returns. Attanasio and Paiella (2011) build a consumption asset pricing model with transaction costs. They jointly estimate the coefficient of relative risk aversion and a lower bound for one-time market participation costs and show that the model can explain the equity premium.

Extending these prior studies, we show that the definition of who the stockholders are is complicated by the entry/exit decisions. In particular, we simulate data from our model and study 3 classifications of stockholders. As in Mankiw and Zeldes (1991), we define "common" stockholders as those that participate in a given period regardless of their participation history. As in Malloy, Moskowitz, and Vissing-Jørgensen (2009), we define "common wealthy" stockholders, the top one third of the wealthiest common stockholders. Lastly, to account for the entry and exit decisions, we define "long-term stockholders" those who own stocks and stay in the stock market for at least 70% of the periods in our sample of simulated data.

We find that long-term stockholders bear the greatest consumption risk (see Figure 1). Also, the consumption growth of long-term stockholders is the most correlated with stock returns. These theoretical findings confirm the empirical evidence in Attanasio, Banks, and Tanner (2002) and Vissing-Jørgensen (2002a) who find that the consumption growth of common stockholders is more volatile and more related to stock returns compared to non-stockholders. We complement their findings and suggest that a more refined stockholder classification is those who own stocks for the long term.

Finally, our work is closely related to Alan (2006) and Bonaparte, Cooper, and Zhu (2012). Alan (2006) studies the decision to own stocks in a model with fixed entry costs.

She also estimates the model with simulated method of moments by matching moments from the 1984 and 1989 waves of the PSID. Bonaparte, Cooper, and Zhu (2012) ignore the decision to participate in the market and instead focus on how consumption smoothing and portfolio rebalancing is affected by portfolio adjustment costs. Compared to these studies, we examine the dynamic decision to enter/exit the market using a more comprehensive portfolio choice model, which is estimated with a more extensive set of moments.

The rest of the paper is organized as follows. Section 2 reports key statistical evidence related to the entry and exit from the stock market. Section 3 presents the model. Section 4 reports the estimation of the model and related findings. Section 5 discusses the implications of our findings for asset pricing tests. Section 6 concludes the discussion.

2 The dynamics of entry and exit: stylized facts

In this section, we illustrate the strong dynamics of entry and exit based on key statistics from the Panel Study of Income Dynamics (PSID) and the Survey of Consumer Finances (SCF). In Appendix D, we report the definitions of our key variables from the PSID and SCF data sets.

The stockholders in our sample are those who own stocks in non-retirement accounts. Specifically, shareholders are those who directly or indirectly own shares of publicly held corporations, mutual funds, or investment trusts. We exclude investments in retirement accounts because there is little trading in these accounts (Ameriks and Zeldes 2004a) and our goal is to examine active stock market participation. Please refer to Appendix A for a summary of the various shareholder definitions in the literature.

2.1 Evidence from the PSID

The PSID reports biennial panel data on household income, wealth, and demographics. The PSID is the longest panel data set reporting the stock market participation status of US households. We focus on the waves from 1999 to 2011 because they include detailed information about stock ownership. We treat the 1999 wave as our baseline year and track households for which we know their participation status in all subsequent waves until the 2011 wave.[8] We report various participation statistics in Table 1.

First, we report the stock market entry and exit by PSID wave of the households that appear in the 1999 survey. Specifically, in column 1 of Panel A, we report the fraction of new stockholders in year $t+2$. That is, the fraction of households that were not stockholders in year t but became stockholders in the following year $t+2$ as a fraction of all households in year t. Based on this statistic, in 2001, 9.7% of households who were not stockholders in 1999 become stockholders in 2001. In other words, if in 1999 we had a 100 households, about 10 of them became new stockholder in 2001. In column 2 of Panel A, we present the respective exit statistic, that is, the fraction of households that own stocks in year t but became new non-stockholders in year $t+2$. In the 2001 wave, this exit statistic was 7.5%. On average, 7.3% (8.7%) new households enter (exit) the stock market between 2 waves, which signifies substantial entry and exit from the stock market.

Next, we examine whether the participation status changes between waves. Specifically, in column 4, we report the fraction of non-stockholders in year t that enter the market in year $t+2$ as fraction of all the non-stockholders in year t. In 2001 for example, we find that about 30.9% of non-stockholders in 1999 enter the market. In column 5, we report the respective statistics related to change of status from stock owner to non-stock owner.

[8]We stop at the 2011 wave since we noticed that in the 2013 wave there are abnormally large stock market exits and lower overall stock market participation than previous waves. Most importantly, we do not notice this change in other data sets such as the 2013 SCF. Therefore, we decided to exclude the 2013 wave from our analysis to avoid biasing our inference.

That is, the fraction of stockholders in year t that exit the market in year $t+2$ as fraction of all the stockholders in year t. In 2001, about 23.9% of stockholders in 1999 exit the market. Overall, there are substantial changes in the participation status between waves: on average, 23.8% households that did not own stocks in one wave enter the market in the next wave, and 28.8% of households that owned stocks exit by the next wave.

2.2 Tracking the 1999 PSID stockholders

To better understand how often households enter and exit the market, we track the behavior of the households who report that they owned stock in the 1999 wave. We call these households baseline stockholders. We report their entry/exit frequencies in subsequent waves in Panel B of Table 1. Specifically, Column 1 reports the fraction of the baseline stockholders that participates in the stock market in any of the following waves. Column 2 reports the fraction of the baseline stockholders who have exited the stock market in a particular year. Column 3 reports the fraction of the baseline stockholders who in year t participated in the stock market in every one of the previous waves. Column 4 reports the fraction of baseline stockholders who re-enter the stock market after exiting in the previous wave. The last row in the table reports the averages of each column.

The statistics in Panel B reveal that many of the baseline stockholders exit from the market permanently. For example, about 23.9% of baseline stockholders report that they do not own stocks in the 2001 wave. Moreover, only 50.8% of them also own stocks after 12 years, while only 32% of them participate in the market in all waves. From those that exit the market, the fraction that re-enters diminishes across waves.

Overall, the statistics in Table 1 reveal that there is substantial turnover in entering and exiting the stock market. There is also persistence in the participation status since many stock owners do not exit the market between two survey waves. It is surprising though, that the number of households who always own stocks in every wave is low.

2.3 Evidence from the SCF

For robustness, we report entry and exit statistics using the most recent panel from the Survey of Consumer Finances (SCF). The SCF is by construction a repeated cross-sectional data set. Therefore, we cannot use all the SCF waves to compute statistics that can capture changes in the participation status of a household. Instead, we only use the panel data that are available for 2007 and 2009.

In Table 2, Panel A, we report market participation rates in 2009 conditional on the participation status in 2007. We find that 37.1% of household do not participate in the stock market during the 2007 and 2009 years, while 46.2% of households participate in both years. Furthermore, about 9.3% of households enter the stock market and own stocks in 2009 but not in 2007. Also, 7.4% of households own stocks in 2007 but decide to exit and not own stocks in 2009. The previous statistics imply that only 73.4% of stockholders in 2007 are also stockholders in 2009.[9] Overall the statistics from the SCF confirm our conclusion from the PSID that there is substantial turnover in the market.

In the SCF, as opposed to the PSID, we have detailed information about the portion of wealth allocated to risky assets. We use these wealth data to infer the economic magnitude of the dollar amount of stock market entries and exits. Specifically, in Panel D of Table 2, we compare the volume bought and sold to the average equity held by stockholders, as well as the average labor earnings of stockholders in 2009. We find that the average amount sold (bought) by exiting (entering) households represents around 32.8% (14.1%) of the

[9] We obtain the 73.4% number as follows: from Panel A of Table 2, we know that only 46.2% of households participated in the stock market in 2007 and 2009, 9.3% of households participated in the stock market in 2009 but not 2007, and 7.4% of households participated in 2007 but not on 2009. This implies that all the households that participated in either the 2007 or 2009 waves are: 7.4% + 9.3% + 46.2% = 62.9%. Because only 46.2% owned stocks in both waves, then it has to be that 73.4% (= 46.2 / 62.9) of stockholders in 2007 are also stockholders in 2009.

average equity held by all stockholders in 2009. The dollar value of exits and entries also constitute 66.6% and 28.7% of 2009 average labor earnings, respectively. Taken together, the turnover in the market comprises a large share of the average equity held, and an even larger share of the average stockholder wages.

2.4 Persistence in non-retirement stock ownership

The previous findings highlight that there is turnover in market participation. But, it is also clear that there is persistence in stock ownership, as well. We examine this persistence by estimating probit models with the PSID and SCF data. Specifically, in Table 3, we report estimates from probit regressions where the dependent variable is a dummy variable related to the current stock market participation status (i.e., 1 if the households own stocks, and 0 otherwise). The key independent variable, labeled "Past participation," indicates the stock market participation status in the previous wave. The other independent variables are household demographics, such as white (binary for race; 1 for white and 0 otherwise), age, male (binary for gender; 1 if male and 0 otherwise), education (years of schooling), income and wealth. Panel A of Table 3 reports summary statistics for the variable in the regressions, and Panel B reports the estimation results. Regression 1 reports results using the SCF panel of 2007-2009. Regressions 2-4 use the PSID data. Regression 2 only uses data for the 2007-2009 period. Regression 3 and 4 use the entire sample from 1999 to 2011.

The results demonstrate that the propensity to participate in the stock market depends on the previous participation status. In the SCF regression, the probit estimate is 0.575. In the PSID regression that uses the same period as the SCF this estimate is similar (i.e., 0.507). Using all the PSID waves, the estimate on the past participation variable is 0.483 without accounting for fixed effects (regression 3) and 0.481 when we account for year fixed effects (regression 4). Also, in untabulated results, we find that these probit estimates imply that on average the probability of owning stocks in a given wave conditional on owning stocks in the past wave is about 0.46. Overall, the probit regressions suggest that there is some persistence in stock ownership.

2.5 Determinants of stock market turnover

Our evidence thus far suggest that there is turnover in stock ownership and in the next section we present a canonical portfolio choice model to examine the factors affecting this turnover. To set the stage for the theoretical analysis, we estimate regressions that directly test the core intuition of portfolio choice models related to stock ownership. As suggested by Guiso, Jappelli, and Terlizzese (1996), portfolio choice models with stochastic income imply that trading in stocks should be driven by the desire to smooth income shocks in combination to facing liquidity constraints. That is, investors that face high income risk and severe liquidity constraints would be forced to enter/exit the market often.

Motivated by this core hypothesis, we estimate cross-sectional regressions where the dependent variables measure the turnover in the stock market. Our main independent variables are income and wealth, which serve as proxies for liquidity constraints (Runkle 1991; Guiso, Jappelli, and Terlizzese 1996). The other important independent variable is the standard deviation of income growth, which is our proxy for income risk (Guiso, Jappelli, and Terlizzese 1996; Heaton and Lucas 2000). The regressions also include demographic control variables (i.e., age, race, gender, education).

For the estimation, we use data from the 1999 to 2011 biennial waves of the PSID. We exclude from the sample households that never participated in the stock market in the 1999 to 2011 period so that we can focus on those households that might have used risky assets as vehicle to smooth income shocks. By excluding the households that never participate, we can also separate the participation decision from the active decision to trade stocks.

Finally, we exclude households with income growth higher than 300% and lower than -70% to ensure that our measure of income risk is not unreasonably high.

We present summary statistics of the variables in the regressions we estimate in Panel A of Table 4. In Panel B of Table 4, we present the estimation results.[10] We report regression results for various measures related to the turnover in market participation over the 1999 to 2011 period. In particular, the dependent variables in regressions 1, 2 and 3 are the log of 1 plus the total number of entries, exits, and entries and/or exists, respectively. In regressions 4, 5, and 6, the dependent variables are dummy variables that take the value of 1 if over the sample period the household had at least one entry, one exit, and one entry or exit, respectively. We estimate regressions 1 to 3 with ordinary least squares and estimate regressions 4 to 6 with a probit estimator.

The regression results confirm the findings in the existing literature that demographic characteristics like age, education and gender affect the entry/exit decisions (Calvet, Campbell, and Sodini 2009; Bilias, Georgarakos, and Haliassos 2010). More importantly, we find that income risk and liquidity constraints affect the decision to enter and exit the stock-market. In terms of income risk, the coefficient estimates of the standard deviation of income growth are positive in all regressions and statistically significant in all regressions but regression 5. In terms of liquidity constraints, the coefficient estimates of wealth are always negative and statistically significant. The estimates related to income are negative but they are less significant.

We also find that the regression results related to entry and exit decisions are symmetric because in our sample the total number of exits and entries are positively correlated. In particular, in untabulated results, we find that the correlation between the total number of entries and the total number of exits is 0.78. This correlation is high because if an investor

[10] Exact definitions of all the variables we use are provided in Appendix D.

exits the market in 2 survey waves, for instance, it has to be the case that he owned stocks in at least 2 other survey waves.

The economic significance implied by the estimates of income risk, wealth, and income is not very high. For instance, the estimates of regression 6 suggest that a one-standard-deviation increase in income risk is associated with an increase in the probability of at least one entry or exit of about 0.04 (representing a 5% increase with respect to the mean probability of 0.82). Also, a one-standard-deviation decrease in wealth is associated with an increase in this probability of about 0.02 (representing a 2% increase with respect to the mean probability of 0.82). The implied economic significant is low, probably because we are using a relative short sample. However, it is comforting that the effects of income risk and liquidity constraints are in line with the intuition of portfolio choice models.

3 Household portfolio choice model with transaction costs

In this section, we present a canonical model of portfolio choice to examine the dynamics of stock market participation. In the model, investors face uninsurable income shocks that they try to smooth using financial assets. However, their trading is limited by short-sale and borrowing constraints as well as trading costs. Developing the theoretical model allows us to assess the importance of factors like the opportunity cost of trading and liquidity constraints, which are typically unobservable or difficult to measure in existing data sets.

3.1 Dynamic optimization problem

Our theoretical set up is based on the work of Bonaparte, Cooper, and Zhu (2012). These authors ignore the decision to participate in the stock market. Instead, they only focus on how households rebalance their equity shares to mitigate income risk. We extend their work and explicitly focus on the decision to enter and exit the stock market.

We assume that we have ex-ante identical investors who make decisions based on the value function v.[11] The value function is the investor's maximum over the options of adjusting or not adjusting his holdings. That is:

$$v(y, s_{-1}, b_{-1}, R_{-1}) = max\{v^{\alpha}(y, s_{-1}, b_{-1}, R_{-1}), v^{n}(y, s_{-1}, b_{-1}, R_{-1})\}, \tag{1}$$

where v^{α} and v^{n} are the value functions of adjusting and not adjusting his asset holdings, respectively.

The arguments of the value function are y, s_{-1}, b_{-1}, and R_{-1}. y is the investor's stochastic income. Income follows a persistent 5-state Markov chain that we estimate using data from the PSID. See Section 3.2.3 and Appendix C for more details.

For simplicity, we assume that the investor has access to one risky and one riskless asset. s_{-1} is his holdings of the risky asset. The return from these holdings is R_{-1}. His holdings of the riskless asset is b_{-1} with return r. Therefore, his total financial wealth at the start of a period is $R_{-1}s_{-1} + rb_{-1}$.

3.1.1 Value function under portfolio adjustment

If the stockholder chooses to adjust his portfolio, then his value function v^{α} is:

$$v^{\alpha}(y, s_{-1}, b_{-1}, R_{-1}) = max_{\ s\geq\bar{s},\ b\geq\bar{b}}\ \ u(con) + \ \ \beta E_{R,y_{+1}|R_{-1},y}\ v(y_{+1}, s, b, R) \tag{2}$$

[11] We acknowledge that investor heterogeneity stemming from demographic differences can create life-cycle considerations in portfolio decisions as shown in Cocco, Gomes, and Maenhout (2005) and Gomes and Michaelides (2005). We abstract from such considerations since we want to examine if a simple canonical portfolio choice model can capture the dynamic decision to own stocks.

Above, $u(con)$ is the utility from non-durable consumption con, which we assume has a CRRA form:[12]

$$u(con) = \frac{1}{1-\gamma} con^{1-\gamma} \tag{3}$$

The consumption level con at any period is given by the following budget constraint:

$$con = R_{-1}s_{-1} + rb_{-1} - s - b + y \times \Psi - F^A - C. \tag{4}$$

In the budget constraint (4), s is the total purchases of risky assets that are bounded below by $\bar{s} = 0$. That is, similar to Heaton and Lucas (1996), we assume that the investor cannot short. We eliminate shorting because most retail investors cannot easily short stocks. The variable b is the household's bond holdings, which is bounded below by $\bar{b}$. While we can allow for borrowing in the model, we find that the best model fit arises with tight borrowing constraints in which the investor is not allowed to borrow, that is, $\bar{b} \geq 0$. Among others, prohibiting borrowing is consistent with Aiyagari (1994), and Gomes and Michaelides (2008). Alan (2006) also imposes short-sale and borrowing constraints in her portfolio choice model.

Consumption con in equation (4) is also affected by three types of transaction costs. First, the function C captures on-going trading costs such as commissions, fees, and other costs related to trading like the bid-ask spread. As in Heaton and Lucas (1996), we assume that these trading costs C are proportional to the change in the value of risky asset holdings. That is, they depend on the difference between s_{-1} and s. Please see Section 3.2. for the exact functional form related to the proportional trading costs C.

[12] Some existing research on household finances (e.g. Gomes and Michaelides (2008)) has moved away from the CRRA utility and towards the recursive utility framework of Epstein and Zin (1989). The main advantage of adopting the recursive utility framework is the separation of the elasticity of intertemporal substitution from risk aversion. However, in models that include portfolio adjustment costs, there is no longer a direct link between the inverse of the EIS and the curvature of the CRRA utility function (Bonaparte, Cooper, and Zhu 2012). Therefore, to keep the model simple, we maintain the CRRA framework.

Second, the variable F^A represents fixed costs to trading, such as the costs of maintaining a trading account or similar vehicle. We include this type of fixed cost since most models in the limited participation literature include a one-time fixed cost to entering the market. To be consistent with this work, we also include fixed-type costs to trading. Nevertheless, our estimation will reveal that F^A is essentially zero.

The third and last component of the trading costs is related to Ψ. Ψ is less than 1 and thus affects consumption through a reduction in labor income. We interpret (1-Ψ) as a per-period time-cost to participation. This time cost includes the cost of information gathering, analysis, trading, and time spent on related taxes for direct stockholders (Dumas and Luciano 1991; Vissing-Jorgensen 2004). Alternatively, (1-Ψ) can be interpreted as a delegation cost for those stockholders holding equity indirectly via mutual funds. Indirect stockholders would incur this cost through annual fund expenses, which are the fees charged by funds for portfolio construction and management (Wermers 2000).

We model the time cost to participation as lost income following Gomes and Michaelides (2005) and Alan (2006). The goal of these studies is to capture limited stock market participation and not the dynamic decision to enter/exit the stock market. Therefore, they set the time cost as a portion of permanent income. Because our focus is on the dynamic decision to own stocks, it is more appropriate to model this time cost as a portion of current income.

In the estimation of the model, we estimate F^A and Ψ. Moreover, we obtain estimates of C from actual trades of retail investors as in Bonaparte, Cooper, and Zhu (2012). We present the estimation details of C is Section 3.2. Additionally, we provide a more detailed description of each trading cost component in Appendix E.

3.1.2 Value function under no portfolio adjustment

If the investor chooses not to rebalance, then his value function v^n is:

$$v^n(y, s_{-1}, b_{-1}, R_{-1}) \quad = \quad max_{b \geq \bar{b}} \;\; u(y + rb_{-1} - b) \quad + \quad \beta E_{R, y_{+1}|R_{-1}, y} \;\; v(y_{-1}, s, b, R) \quad (5)$$

In this case, the stockholder consumes only his labor income plus the cash payouts of his riskless asset holdings. For simplicity, we assume that the risky asset return is entirely based on capital gains (no cash dividends paid out). When there is no rebalancing, the proceeds from the existing stock portfolio are costlessly reinvested.[13] Hence,

$$s = R_{-1}s_{-1} \quad (6)$$

Overall, our model is an incomplete market model with restrictions to trading. In the presence of transaction costs as well as short-sale and borrowing constraints, investors in the model cannot fully insulate their consumption from negative income shocks. These limits to trading will sometimes force investors to sell all their stock holdings in response to severe income shocks.

3.2 Proportional trading costs, asset returns, and income process

To close the model, we present the functional forms related to variable trading costs, household income, and asset returns. To aid the estimation of the model, we directly estimate some of the parameters in these functions, rather than estimate them within the dynamic programming model. In the discussion below, we present the parameter estimates.

[13] In our baseline model, we assume that any capital gains are converted into new shares since the price of a share is kept fixed at unity. In an alternative specification that we explore, we assume that the actual shares remain constant, allowing consumption to absorb the return on the existing portfolio. In unreported results, we find that the latter alternative approach does not affect the main conclusions of the model.

For completeness, we report the definitions of all parameters and functions of the model in Appendix F.

3.2.1 Proportional trading costs

The definition of the proportional trading cost function C follows Heaton and Lucas (1996), Vissing-Jørgensen (2002b) and Bonaparte, Cooper, and Zhu (2012). In the model, there is only one risky asset. However, in the data set we use to estimate C, investors trade multiple assets. Therefore, the specification of C that we estimate is based on multiple assets. After the estimation, we adopt the multiple-asset specification to a single-asset specification.

In the estimation, we assume that the cost function C depends on the change in asset holdings of each asset i. That is, C is a function of the differences $s^i_{-1} - s^i$. For simplicity, C is separable across assets and it differs between sales and purchases:

$$C = \sum_i C^j(s^i_{-1}, s^i), \tag{7}$$

where $j = b$ for assets i being bought and $j = s$ for assets being sold. When the stockholder buys asset i, that is $s^i \geq s^i_{-1}$, the functional form for the proportional trading costs follows a quadratic specification:

$$C^b(s^i_{-1}, s^i) = v^b_0 + v^b_1(s^i_{-1} - s^i) + v^b_2(s^i_{-1} - s^i)^2. \tag{8}$$

Similarly, when the stockholder sells asset i, the functional form for the proportional trading costs is as follows:

$$C^s(s^i_{-1}, s^i) = v^s_0 + v^s_1(s^i_{-1} - s^i) + v^s_2(s^i_{-1} - s^i)^2. \tag{9}$$

To ensure that the trading costs captured by C are consistent with what investors face, we adopt the approach of Bonaparte, Cooper, and Zhu (2012) and directly estimate equations (7) and (8) with monthly stockholder trading data. Specifically, we use the Barber and Odean (2000) data and focus on trades of common stocks. The data contains information on common stock trades of about 78,000 stockholders who were clients at a discount brokerage firm from January 1991 to December 1996.

In our sample, we have over 3 million observations where in each observation a stockholder (trader) reports: trade date, buy or sell, quantity of shares transacted, commission (in dollar value), CUSIP identifier and the price. If a stockholder bought different stocks in a given month, the stockholder reports the commission, quantity, and price for each one of these stocks separately. Based on this data, we compute household trading costs. They include direct costs such as brokerage frees and commissions as well as opportunity costs of trading from unfilled or partially filled limit orders. Moreover, they account for the bid-ask spread.

We estimate the trading cost equations (7) and (8) with ordinary least squares (OLS). In the OLS regressions, the dependent variable is the transaction costs. The independent variables are the trade value (i.e., price of the shares times the quantity of shares), the trade value squared, and a constant. We report the results in Appendix B.

The estimation suggests that the average cost of trading, captured by the constant in the regressions, is about \$56 for buying and \$61 for selling. The estimates of the linear and quadratic terms are also important. To get a sense of magnitudes, the average purchase (sale) in our sample has a value of about \$11,000 (\$13,372), and thus, the cost of this trade is about \$70.00 (\$80). For trades of this size, the impact of the quadratic term is small.

We acknowledge that these trading costs might appear high since they are estimated using data from the 90's. Since then, trading costs have been declining (Bogan 2008).

Nevertheless, as we show in Section 4.6, the variable trading costs do not drive the results of the model. We include them in model because investors are still exposed to the bid-ask spread even if the commissions and fees for trading are low.

3.2.2 Asset return processes

In the model we allow for two assets: a riskless asset and a risky asset. The return of the riskless asset is from Bonaparte, Cooper, and Zhu (2012). Specifically, we set it equal to 1.0% annually (0.25% quarterly).

The risky asset represents the stock market return and it follows an IID process with 2 states.[14] This assumption is consistent with Bonaparte, Cooper, and Zhu (2012) who find that the estimated serial correlation of annual and quarterly returns is not significantly different from zero.

In our full model estimation, we estimate the average market return R within the model. For simplicity however, we assume that its quarterly standard deviation is 8.3%. We obtain the standard deviation of the market return using the real returns including dividends of the S&P500 index from the web site of Robert Shiller for the period 1947-2007.[15] The inclusion of dividends in the stock market return is consistent with our model of inaction where dividends are costlessly reinvested.

3.2.3 Income process

The literature has found that the estimation of household-level income processes is quite difficult. For example, see Guvenen (2007) and Browning, Ejrnaes, and Alvarez (2010). Therefore, we have chosen a simple model for income that the previous literature has shown

[14] In unreported results we find that adding more return states does not change the results significantly. Therefore, to keep the model simple, we only include 2 states in the main analysis.

[15] These data are available at http://www.econ.yale.edu/ shiller/data.html. We choose the 1947-2007 period since it is similar to the sample periods of the other data sets that we use.

can reasonably capture the evolution of household income (Viceira 2001; Campbell, Cocco, Gomes, and Maenhout 2001; Gourinchas and Parker 2002; Gomes and Michaelides 2005; Gomes and Michaelides 2008).

In particular, we decompose income into two main components. The first one represents the deterministic component of income and depends on demographic variables such as age and education. The second component is the stochastic part of income. We assume that the stochastic part follows an AR(1) process and it is affected by an idiosyncratic income shock. We estimate the income model using data from the PSID. We provide more details about the estimation in Appendix C.[16] After the estimation, we transform the income process into a 5-state Markov chain using the methodology of Tauchen (1986).

4 Simulated method of moments estimation

Conditional on the estimates of the processes related to the proportional trading costs, income, and returns, we estimate the remaining model parameters with simulated method of moments. These parameters are γ, Ψ, β, F^A, and R. We use 13 data moments to identify these deep parameters. We use moments related to household-level decisions and aggregate-level moments. We include aggregate moments to ensure that the aggregate implications of the model are consistent with aggregate-level stylized facts.

4.1 Entry, exit and stock-market participation moments

Our first household-level data moments are related to the entry and exit decisions. Specifically, we select moments that can capture the entry and exit decision in the short-

[16] As we explain in the Appendix C, the empirical process for income y of household i at time t is $y_{i,t} = \tau Z_{i,t} + A_{i,t}$. Z includes the demographic variables. A is the persistent component of income, $A_{i,t} = \rho A_{i,t-1} + \epsilon_{i,t}$, and ϵ_t is the transitory shock. Among others, Campbell, Cocco, Gomes, and Maenhout (2001) use a very similar income process in the calibration of their model.

term (2 years) and long-term (12 years).

The first moment is related to the probability of participating in the market today conditional on having participated in the recent past. We base this moment on the estimates of the variable "past participation" from the probit regressions reported in Panel B of Table 3. In the SMM estimation, we set this probit estimate to 0.51, which is the average across the 4 estimates in Table 3, Panel B. The second moment is related to the average entry and exit rates. Based on the PSID statistics in Panel A of Table 1, we assume that both of these rates are 8%. The 8% number is the average of the mean entry and exit rates between 2 consecutive waves, which are 7.3% and 8.7%, respectively. These two moments help us to capture the short-term entry and exit turnover.

The third moment is related to the portion of households that always participate. We set this portion to 32.8%, which is the portion of stockholders in the PSID that participated in all 12 waves (see Table 1, Panel B, Column 3). The fourth moment is also taken from the PSID and it is the portion of households (50.8%) that own stocks in 1999 and 2011 (see Table 1, Panel B, Column 1). These two moments help us capture the long term entry and exit turnover.

Another moment we use is related to the rebalancing rate of the risky equity share. We define the rebalancing rate as the cross-sectional average of our trading indicator. The trading indicator takes the value of 1 if the household has changed its asset holdings in a given period and zero otherwise. We use the estimate of the rebalancing rate from the PSID. Specifically, in Table 1, Panel C, we show that the average rebalancing rate is 48.6%. This moment is important because it can help identify the per-period time-costs of participation captured by (1-Ψ), as well as the fixed-type costs of participation captured by F^A.

The fourth moment is related to average stock market participation. We set this rate to

48.9%, which is the average stock market participation in 6 SCF waves (see Table 2, Panel A). We use this statistic because it is in line with Grinblatt, Keloharju, and Linnainmaa (2011) who document that about 50% of US households own stocks.

4.2 Wealth ratio moments

We use three moments related to wealth ratios. For the stockholders, we use the ratio of median financial wealth to median wages and their median stockholdings to median wealth. For all households (i.e., stockholder and non-stockholders), we use the ratio of median financial wealth to median wages. We obtain these statistics from the 6 SCF waves from 1989 to 2010. Statistics for stockholders are in Table 2, Panel E.1, and statistics for all households are in Panel E.2. Based on these statistics, we set the stockholders' average equity share to 0.41 and their average wealth-to-wages ratio to 1.03. In the case of all households, we set the average wealth-to-wages ratio to 0.42. We use the SCF for these moments because the SCF wealth information is more precise than that in the PSID.

4.3 Consumption-market return sensitivity and risk premium moments

The final set of moments is related to market-level moments. Three of these moments capture the sensitivity of consumption growth to stock market returns. To obtain these moments, we estimate the log-linear approximation of a consumption Euler equation, as in Hansen and Singleton (1983). Specifically, we estimate OLS regressions of the log of consumption growth on the log market return:

$$\log(\frac{c_{t+1}}{c_t}) = \alpha_0 + \alpha_1 \times \log(R_{t+1}) \tag{10}$$

We offer no structural interpretation of the parameter α_1. We interpret it as an empirical moment and call it the aggregate EIS to distinguish it from the EIS at the stockholder level. We estimate (10) for different stockholder consumption growth horizons, namely 1, 4 and 16 quarter horizons. The motivation for using different lengths of quarterly horizons is to examine if the impact of trading costs is less important in the long run.

We estimate the consumption growth regressions using the non-durable consumption data of stockholders created by Malloy, Moskowitz, and Vissing-Jørgensen (2009). The data are quarterly and cover the period from March 1982 to November 2004. For the return data, we use the monthly S&P returns from the web site of Robert Shiller. We convert the monthly return data to quarterly in order to match the quarterly frequency of the consumption data. In unreported results we find that the aggregate EIS for time horizons of 1, 4 and 16 quarters are 0.0295, 0.1269 and 0.3518, respectively. We label these estimates as "$\alpha_1 - 1$ quarter," "$\alpha_1 - 8$ quarter," and "$\alpha_1 - 16$ quarter," respectively.

The final moment is related to the risk premium. We compute the risk premium using the stock market data from the web site of Robert Shiller for the period 1947-2007. In this period, the risk premium was about 7.6% on an annualized basis.

4.4 Simulated method of moments estimation set up

We estimate the model with Simulated Methods of Moments (SMM). The SMM is based on the following minimization problem:

$$\prod = min_{(\gamma,\Psi,\beta,F^A,R)}(M^s - M^d)'W(M^s - M^d) \tag{11}$$

Above, M^d are the moments from the data and M^s are the simulated values of those moments for a given set of parameters. The matrix W is an identity matrix. Unfortunately, we cannot use the optimal weighting matrix, which is usually the Newey-West (1986) covariance matrix of the data moments. This is case because the moments in M^d are from two data sets, the SCF and the PSID. Thus, it is not possible to calculate covariances of moments from different sources that also refer to different types of households. In particular, the moments related to the stock shares from SCF are based on a subsample containing only stockholders. In contrast, the participation rates and wealth/income ratios are based on an entire sample that include non-participants as well as participants.

To apply the simulated method of moments, we solve the dynamic programming problem of an individual household. We solve the model with value function iteration. We then create a panel data set with 500 households and 800 quarters. We compute moments from the simulated data set in exactly the same manner that the moments were calculated in the actual data. The estimation finds the parameters $(\gamma, \Psi, \beta, F^A, R)$ that bring these simulated moments as close as possible to the actual data moments. See Appendix G for more details on the simulation.

4.5 SMM baseline estimation results

We report in Table 5 our main estimation results. In Panel A, we report the estimates of the deep parameters of the model. In Panel B, we report the data moments and the respective moments implied by the model.

4.5.1 Parameter estimates

As we see in Panel A, the estimate of the coefficient of relative risk aversion is 3.176. This estimate is quite reasonable and in line with the existing literature. For example, Cagetti (2003) matches wealth data and reports estimates of the coefficient of relative risk aversion around 2.74 to 4.26. As Alan (2006) notes, estimates based on matching consumption data report lower estimates for the coefficient of risk aversion. For example, in Attanasio, Banks, Meghir, and Weber (1999) the estimate of the coefficient of relative risk aversion is about 1.5 while in Gourinchas and Parker (2002), it is between 0.28 and

2.29.

The estimate for the discount factor is 0.963. This estimate is also in line with existing findings. In particular, Alan (2006) finds the discount factor to be about 0.86 while Gourinchas and Parker (2002) report an estimate of about 0.96.

The estimated parameters for the trading costs are also reasonable. The fixed participation cost F^A is 0.020 units of consumption, representing a negligible portion. This is not surprising since most of the literature finds that fixed-type costs to stock ownership are very low. For instance, using a different methodology, Paiella (2007) finds that the lower bound of fixed costs that can rationalize not participating in the market is $130 per year. Attanasio and Paiella (2011) estimate these costs to be only $72 per year.

The per-period time-cost of participation is about 3.20% of income (i.e., $0.032 = (1 - 0.968)$ since $\Psi = 0.968$). At $\Psi = 0.968$, the cost of portfolio rebalancing is about $2,304 when using the average annual labor income from our PSID sample (i.e., $72,000) . This estimate is fairly reasonable especially when compared to the actual costs of portfolio management delegation. Specifically, Wermers (2000) finds that the average expense ratio of active mutual funds (weighted by total net assets of funds) is around 0.93% over the 1990-1994 period.[17]

We use the 0.93% expense percentage and the wealth of the average household invested in equity to come up with a back-of-the-envelope estimate of the costs of delegation. Based on the SCF wealth statistics in Panel B of Table 2, we find that the average implied delegation cost ranges between $1,485 and $2,154 ($1,471 and $2,019) for the stockholders in 2007 (2009).[18] These approximate delegation costs are close to our time costs estimate

[17]We confirm that this estimate is similar to more recent figures using industry reports. Please refer to https://www.kitces.com/blog/independent-financial-advisor-fees-comparison-typical-aum-wealth-management-fee/.

[18]For example, from Table 2, Panel C, the mean equity holdings of stockholders in 2007 is $231,641, for which our back-of-the-envelope calculation finds a cost of $2,154 = (0.93% x $231,641).

of \$2,304.

Furthermore, the model implies an equity premium that is close to the data. This result is consistent with Heaton and Lucas (1996). Their model includes idiosyncratic income shocks and borrowing constraints. Because in their model markets are incomplete, equilibrium consumption growth across investors is not equalized, which leads to a sizeable equity premium. Attanasio and Paiella (2011) confirm that a model with transaction costs and limited stock-market participation can generate a reasonable equity premium.

In the estimation, we also compute the precision of our parameter estimates. In particular, we estimate the standard errors following the methods of Adda and Cooper (2003). Based on these standard errors, we find that all our parameter estimates are statistically significant.

4.5.2 Model fit and implied moments

We also assess the empirical fit of the model with the J-test of the over-identifying restrictions. We find that the J-statistic is 0.104. With eight degree of freedom, the respective p-value is about 0.010. Thus, the model is statistically not rejected.

The high fit of the model suggests that the empirical models are close to those implied by the model. As we see in Panel B of Table 5, the model captures the average participation rate relatively well, which is 0.492 in the data and 0.613 in the model. Moreover, the rebalancing rate in the model (0.599) is relatively close to the one in the data (0.486). Also, the probit estimate related to the persistence in participation between two periods is similar in the model (0.443) and in the data (0.510).

The model predicts slightly higher turnover in market participation relative to the data. For instance, the average entry/exit in the model is 0.286 but in the data is only 0.160. This finding is not surprising since in the model, households can only trade one risky asset to mitigate income and return shocks. Also, the median relative financial wealth to wages for stockholders is higher in the data (1.034) than in the model (0.886). Finally, as expected, upon participation in the stock market, the households in the model invest more in risky assets than in the data (average equity share in the data is 0.413 and 0.784 in the model).

Overall, our model fits the data moments relatively well with preference parameters that are reasonable.[19] Next, we use the simulation of our model to examine the key determinants of the stock-market entry and exit decisions.

4.6 Transaction costs and participation decisions

In this section, we examine which type of transaction cost is the most important for the dynamic decision to own stocks. We report how the model-implied moments respond to changes in one or more of the transaction cost components. Specifically, we use the estimates of the preference parameters from Table 5 and simulate the model by shutting off some components of the transaction cost function. Then, we report in Table 6 the implied moments from these simulations.

The results in Table 6 show that transaction costs are important but not all cost components are equally significant. Specifically, we see in column 3 that when we set all transaction costs to zero, the model predicts excess participation compared to the data. Therefore, transaction costs are needed to fit the participation turnover.

Next, we introduce the various forms of transaction costs. The introduction of fixed-type costs to participation, column 4, only marginally improves the predictions of the model. The model generated moments get closer to the data moments in column 5 when we allow for both fixed-type and per-period time-costs to participation. The implied model moments again diverge from the data moments in column 6 when we eliminate the per-

[19] Please see Appendix H for the sensitivity analysis related to our main estimation results.

period time-costs and allow for fixed-type along with proportional trading costs.

Overall, any kind of monetary costs (either fixed or proportional trading costs) do not seem to matter for the decision to participate in the market. This finding is consistent with the evidence in the literature (Gomes and Michaelides 2005). We find that what matters the most is per-period time cost to participation.

4.7 Income risk and participation decisions

Next, we use our simulation results to examine the importance of income and income risk for the endogenous decision to enter/exit the market. For this analysis, we use simulated data from the model and estimate cross-sectional regressions of entry/exit moments. In the regressions, the main dependent variables are the fraction of times a household enters or exits. Another dependent variable is the fraction of periods a household participates in the stock market. The independent variables are the realized standard deviation of household income growth and the average household income level.

We report the cross-sectional household regressions results in Table 7. The results demonstrate that the income risk and income level influence the entry, exit, and overall participation in the stock market. Households with higher income risk have higher frequency of entry and exit and have higher participation in stock market. In contrast, households with higher average income level tend to have lower frequency of entry and exit and higher participation in the stock market. These results are consistent with the estimates in Panel B of Table 4 showing that PSID households with lower wealth and income that face higher income risk had more entries and exits from the stock market.

These findings are intuitive. To begin with, if one has high income risk, he needs to participate in the market to smooth income shocks. As Carroll (2001) explains, households with higher income risk need to buy "buffer stocks" to mitigate consumption risk and thus have a high participation rate. But, these high-income risk households also need to sell off their stock holdings in bad times to fund consumption. Thus, they have high entry and exit rates as well as higher overall stock-market participation. In contrast, household with high income can self-insure, thus they participate more and do not need to trade too much to mitigate the impact of income shocks.

To further understand the impact of income risk, we run a simulation exercise where we increase income risk. Specifically, we magnify the impact of idiosyncratic income shocks on total income by 25%.[20] In Table 8, we report the new model-implied moments. We find that the average stock market participation increases to 73% compared to 61.3% from the baseline simulation. Also, the rebalancing rate of the equity share increases to 71.9% from 59.9%. Overall, consistent with the findings from the cross-sectional regressions in Table 7, we find that high income risk induces more participation and more rebalancing in order to mitigate the impact of income risk on consumption.

4.8 Aggregate shocks and the dynamics of stock market participation

In our previous analysis, we focused on how household-level factors affect the decision to enter/exit the stock market. Next, we rotate the point of interest and focus on aggregate-level shocks. At the aggregate level, our model allows shocks related to market returns and the aggregate component of income. To examine the impact of these aggregate shocks, we aggregate the household-level simulated data in each period and estimate time-series regressions.

Specifically, we first define average entry (exit) as the fraction of households who enter (exit) the stock market in a given period. We define average participation as the fraction

[20]Note that the empirical process for income y from Appendix C is $y_{i,t} = \tau Z_{i,t} + A_{i,t}$. A is the persistent component of income, $A_{i,t} = \rho A_{i,t-1} + \epsilon_{i,t}$, and $\epsilon_{i,t}$ is the transitory shock. In our income risk amplification exercise, we multiply $\epsilon_{i,t}$ by 1.25 in the process for A, i.e., $A_{i,t} = \rho A_{i,t-1} + 1.25 \times \epsilon_{i,t}$. Then, we obtain a new 5-state Markov chain that we use in the new simulation.

of households who participate in the stock market in a given period. Also, we define total flows as the net cash flow in the stock market, i.e., total purchases of stocks minus total sales of stocks. Then, we estimate regressions to measure the impact of stock returns and average (across all households) income levels on the four aforementioned variables. We report the regressions results in Table 8. Since we calculate income growth rates, the length of our regression sample is 799 observations.

We find that the stock market return strongly influences the average entry and exit from the stock market, the average participation, as well as the total flow. These effects are all statistically significant. Average income growth is also important. The estimates show that an increase in average income growth leads to higher average entry, lower average exit, lower average participation, and high flow of capital in the stock market. These income-growth effects are all statistically strong except from the relationship between income growth and average participation.

The results of the time-series regressions are quite intuitive. They suggest that in periods when the market is doing well, households want to invest more in the stock market to increase their lifetime wealth. Also, in periods when aggregate income is rising, households have more liquidity to invest. But, in these periods households might not have a strong incentive to participate because with overall higher income growth they do not need to invest in "buffer" stocks to smooth consumption.

Overall, the previous analysis shows that a canonical portfolio choice model can capture relatively well the decision to enter/exit the stock market. Specifically, the estimates for deep model parameters are consistent with the literature. The model also fits the data moments well and cannot be rejected by the J-test of over-identifying restrictions. Our analysis also points out that the important trading costs are those related to income lost due to time spent on portfolio rebalancing. Moreover, household income risk as well as aggregate income and stock market shocks affect the dynamic decision to participate in the stock market.

5 Asset pricing implications

In this section, we examine potential implications of the large turnover in stock market participation for asset pricing tests. In particular, the entry/exit from the market can complicate the definition of who the marginal investors are, and in extension, the measurement of the relevant sources of stockholder risk. The sources of risk affecting stockholders are the key quantities in asset pricing models that strive to explain the equity premium and the cross-section of expected returns.

We illustrate why the entry/exit from the market can affect the inference of consumption CAPMs by examining whether the participation history of investors affects the properties of the empirical moments related to stockholders. In particular, we follow Attanasio, Banks, and Tanner (2002) and Vissing-Jørgensen (2002a). Extending the work of Mankiw and Zeldes (1991), these authors use household-level data from the U.K. and the U.S. and show that consumption growth risk of market participants is higher than non-participants. They also show that the consumption growth of asset holders is more correlated with market returns than the consumption growth of non-holders. Therefore, both studies confirm the findings in Mankiw and Zeldes (1991) who argue that consumption CAPMs should be tested using data for stock-holders.

5.1 Stockholders classifications conditional on participation history

We examine three stockholder classifications. In the first classification, we follow Mankiw and Zeldes (1991) and classify as stockholders those who own stocks at a given period, regardless of their ownership status in other periods. We call these stockholders "common" stockholders. In the second classification, we follow Malloy, Moskowitz, and Vissing-Jørgensen (2009), and focus on "common wealthy" stockholders. The common wealthy stockholders are the top-third wealthiest common stockholders in each period.

In the third classification, we explicitly take into account the entry/exit decisions. That is, we focus on the common stockholders that also participate in other periods. We call them "long-term" stockholders and they are the common stockholders whose record shows that they owned stocks in more than 70% of the time periods in our sample.

Using the data from our simulation exercise, we first report in Panel A of Table 10 various wealth and income statistics for these 3 groups. We find that the common stockholders have the lowest wealth relative to average wealth. They also have the lowest income relative to average income and the lowest income risk. In contrast, the long-term stockholders and the wealthy stockholders are relatively wealthier, with higher income but also higher income risk. These statistics imply that because of higher wealth, the long-term stockholders and the wealthy stockholders can withstand more negative income and return shocks.

5.2 Participation history and consumption risk

Next, we compute the consumption risk of the different stockholder groups over different horizons. Specifically, we first compute the consumption growth in a given period for each household. Then, in each period, we take a cross-sectional average among the households in each stockholder group to create a time series of the average consumption growth by stockholder group. Finally, we calculate the standard deviation of these time series over various horizons to capture consumption risk. Similar to Malloy, Moskowitz, and Vissing-Jørgensen (2009), we examine 1, 4 and 16 quarter horizons and present the consumption risk estimates in Figure 1.

As we see in Figure 1, the consumption risk of the long-term stockholders (i.e., those that frequently participate in the stock market) is larger than that of the common or common wealthy stockholders. This finding implies that high turnover in and out of the stock market can ease consumption risk. However, long-term stock holders have relatively high wealth and income. Being wealthier, they can withstand negative income shocks and they do not need to completely sell their risky asset holdings when faced with such negative shocks.[21]

This is a very interesting finding since most asset pricing models cannot explain the equity premium with reasonable levels of risk aversion because observed consumption risk is usually too low. Our results demonstrate that consumption risk depends on the frequency of entry and exit. Consumption risk is the highest for those households that have the lowest exit frequency. Thus, they should be rewarded with a high risk premium even if they have low relative risk aversion.

5.3 Participation history and estimates of the EIS

The evidence in Figure 1 suggests that long-term stockholders (those who almost always participate in the market) bear greater consumption risk. Thus, empirical tests of consumption asset pricing models that use the consumption data of long-term stockholders may be more successful in pricing expected returns than using the consumption data of "common" stockholders or "common wealthy" stockholders. This will be the case if the consumption growth of long-term stockholders *also* has the highest correlation with market returns.

We follow Attanasio, Banks, and Tanner (2002) and estimate the relationship between

[21] In untabulated results, we find that by staying in the market longer, investors can enjoy the full equity premium. Therefore, if an investor can avoid exiting the market due to a negative income shock, he will rationally do so to improve long-term performance and increase life-time wealth.

consumption growth and market returns. Specifically, we use the log linearized Euler equation of Hansen and Singleton (1983) for market participants and obtain an estimate for the aggregate EIS. This estimation approach is similar to that of Vissing-Jørgensen (2002a) who suggests that Euler equation estimation is only valid for those how own financial assets. To obtain the estimated EIS, we focus on the 16-quarter time horizon and use data from the simulation of our model. We use 16-quarter to be consistent with Malloy, Moskowitz, and Vissing-Jørgensen (2009).

For each of our 3 stockholder classifications, we estimate different OLS time-series regressions. The dependent variable is the average 16-quarter consumption growth and the dependent variable is the 16-quarter market return. The time-series for average consumption growth is based on the cross-sectional average of the 16-quarter consumption growth of all households in each stockholder classification. The 16-quarter return is the exponential sum of the market return for 16 quarters.

We report the results in the last row of Table 10. Our estimation results show that the estimates of the EIS depend on the stockholder classification. Most importantly, we find that the EIS obtained from the sample of the long-term stockholder is the highest and is 0.236. This estimated EIS is lower for “common wealthy” and “common” stockholders and is equal to 0.176 and 0.111, respectively.

5.4 Intuition: Marginal investors and long-term stockholders

We obtain the highest EIS from the sample of long-term stockholders because they are most responsive to stock market return shocks and they rarely exit from the market. Specifically, we find that their rebalancing rate (see Table 10) is very high and it is around 71.7%. Because they trade a lot and they rarely completely exit the market, they are almost always a significant portion of the group of marginal investors.

In contrast, the rebalancing rate of the common and common wealthy stockholders is lower. The lowest rate is for the common stockholder (32.3%). These stockholders enter and exit the market frequently and they are not consistently a part of the marginal investor group. Therefore, their consumption growth data are not strongly related to market returns, resulting in a low EIS.

To summarize, the way we classify stockholders can affect asset pricing tests. We show that in an economy with trading frictions, borrowing constraints, and income shocks, the households that stay in the market the longest are those that end up having high consumption risk. Moreover, since their consumption growth exhibits the highest comovement with market returns, they should be rewarded with a high equity premium.

6 Conclusion

In this paper, we examine the dynamic decision to own stocks. Our analysis is motivated by the stylized fact that when we exclude investments in retirement accounts, about 25% of stockholders enter or exit the stock market at the biennial frequency. To examine the dynamic decision to enter and exit the stock market, we estimate a model that includes transactions costs, borrowing constraints, short-sale constraints and income shocks. The estimation is done with the simulated method of moments estimator.

Our estimation shows that our model canonical portfolio choice model can fit many moments related to the decision to enter/exit the market. Simulated data from the model suggest that households with high frequency of entry and exit are exposed to high income shocks. Moreover, among the various transaction costs that we consider, we find that per-period time cost to participation is the most important in fitting the entry/exit moments.

More broadly, our results contribute to the asset pricing literature that focuses on the behavior of stockholders. We show that key moments of consumption growth depend on the classification of who the stockholders are. Specifically, we find that long-term stockholders exhibit high consumption growth risk and their consumption growth is highly correlated with market returns. Therefore, our evidence suggests that empirical research should focus on the behavior of long-term investors when studying asset pricing phenomena.

In future work, it will be important to study a more general model that includes multiple risky assets. One such risk asset can be a retirement account. Like existing retirement accounts, it can offer tax benefits and an alternative way to save. This expanded model can help us understand why investors rarely rebalance their retirement accounts while, as we show, they seem to trade a lot when it comes to non-retirement accounts.

References

Adda, J. and R. W. Cooper (2003). *Dynamic economics: quantitative methods and applications*. MIT press.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics 109*(3), 659–684.

Alan, S. (2006). Entry costs and stock market participation over the life cycle. *Review of Economic Dynamics 9*, 588–611.

Ameriks, J. and S. P. Zeldes (2004a). How do household portfolio shares vary with age. Technical report, working paper, Columbia University.

Ameriks, J. and S. P. Zeldes (2004b). How do household portfolio shares vary with age? *Working Paper, Columbia University, Graduate School of Business*.

Angerer, X. and P.-S. Lam (2009). Income risk and portfolio choice: An empirical study. *The Journal of Finance 64*(2), 1037–1055.

Attanasio, O. P., J. Banks, C. Meghir, and G. Weber (1999). Humps and bumps in lifetime consumption. *Journal of Business & Economic Statistics 17*(1), 22–35.

Attanasio, O. P., J. Banks, and S. Tanner (2002). Asset Holding and Consumption Volatility. *Journal of Political Economy 110*(4), 771–792.

Attanasio, O. P. and M. Paiella (2011). Intertemporal Consumption Choices, Transaction Costs and Limited Participation in Financial Markets: Reconciling Data and Theory. *Journal of Applied Econometrics 26*, 322–343.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance 55*(2), 773–806.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics 116*(1), 261–292.

Barrot, J.-N., R. Kaniel, and D. Sraer (2016). Are retail traders compensated for providing liquidity? *Journal of Financial Economics 120*(1), 146–168.

Bertaut, C. C. (1998). Stockholding behavior of us households: Evidence from the 1983–1989 survey of consumer finances. *Review of Economics and Statistics 80*(2), 263–275.

Bertaut, C. C. and M. Haliassos (1997). Precautionary portfolio behavior from a life-cycle perspective. *Journal of Economic Dynamics and Control 21*(8-9), 1511–1542.

Betermier, S., L. E. Calvet, and P. Sodini (2017). Who are the value and growth investors? *The Journal of Finance 72*(1), 5–46.

Betermier, S., T. Jansson, C. Parlour, and J. Walden (2012). Hedging labor income risk. *Journal of Financial Economics 105*, 622–639.

Bilias, Y., D. Georgarakos, and M. Haliassos (2010). Portfolio inertia and stock market fluctuations. *Journal of Money, Credit and Banking 42*(4), 715–742.

Bogan, V. (2008). Stock Market Participation and the Internet. *Journal of Financial & Quantitative Analysis 43*(1), 191–211.

Bonaparte, Y., R. Cooper, and G. Zhu (2012). Consumption smoothing and portfolio rebalancing: The effects of adjustment costs. *Journal of Monetary Economics 59*(8), 751–768.

Bonaparte, Y., G. Korniotis, and A. Kumar (2014). Income hedging and portfolio decisions. *Journal of Financial Economics 113*, 300–324.

Brav, A., G. M. Constantinides, and C. C. Geczy (2002). Asset pricing with heterogeneous consumers and limited participation: Empirical evidence. *Journal of Political Economy 110*(4), 793–824.

Browning, M., M. Ejrnaes, and J. Alvarez (2010). Modelling income processes with lots of heterogeneity. *The Review of Economic Studies 77*(4), 1353–1381.

Cagetti, M. (2003). Wealth accumulation over the life cycle and precautionary savings. *Journal of Business & Economic Statistics 21*(3), 339–353.

Calvet, L., M. Gonzalez-Eiras, and P. Sodini (2004). Financial innovation, market participation, and asset prices. *Journal of Financial and Quantitative Analysis 39*(3), 431–459.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy 115*, 707–747.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2009). Fight or flight? portfolio rebalancing by individual investors. *The Quarterly journal of economics 124*(1), 301–348.

Campbell, J. Y. (2006). Household finance. *Journal of Finance 61*, 1553–1604.

Campbell, J. Y., J. F. Cocco, F. J. Gomes, and P. J. Maenhout (2001). Investing retirement wealth: A life-cycle model. In *Risk aspects of investment-based Social Security reform*, pp. 439–482. University of Chicago Press.

Campbell, J. Y. and L. M. Viceira (2002). *Strategic asset allocation: portfolio choice for long-term investors*. Oxford University Press, USA.

Carroll, C. D. (2001). Death to the log-linearized consumption euler equation!(and very poor health to the second-order approximation). *Advances in Macroeconomics 1*(1).

Carroll, C. D. and A. A. Samwick (1997). The nature of precautionary wealth. *Journal of monetary Economics 40*(1), 41–71.

Cocco, J. F., F. J. Gomes, and P. J. Maenhout (2005). Consumption and portfolio choice over the life cycle. *Review of financial Studies 18*(2), 491–533.

Constantinides, G. M. (1976). Stochastic cash management with fixed and proportional transaction costs. *Management Science 22*(12), 1320–1331.

Constantinides, G. M. (1986). Capital market equilibrium with transaction costs. *Journal of political Economy 94*(4), 842–862.

Dorn, D. and M. Weber (2017). Losing Trust in Money Doctors. *Working Paper*.

Dumas, B. and E. Luciano (1991). An exact solution to a dynamic portfolio choice problem under transactions costs. *The Journal of Finance 46*(2), 577–595.

Epstein, L. G. and S. E. Zin (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica 57*(4), 937–969.

Gârleanu, N. and L. H. Pedersen (2013). Dynamic trading with predictable returns and transaction costs. *The Journal of Finance 68*(6), 2309–2340.

Gennotte, G. and A. Jung (1994). Investment strategies under transaction costs: the finite horizon case. *Management Science 40*(3), 385–404.

Gollier, C. (2001). Wealth inequality and asset pricing. *The Review of Economic Studies 68*(1), 181–203.

Gomes, F. and A. Michaelides (2005). Optimal life-cycle asset allocation: Understanding the empirical evidence. *The Journal of Finance 60*(2), 869–904.

Gomes, F. and A. Michaelides (2008). Asset pricing with limited risk sharing and heterogeneous agents. *Review of Financial Studies 21*(1), 415–448.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the life cycle. *Econometrica 70*(1), 47–89.

Grinblatt, M., M. Keloharju, and J. Linnainmaa (2011). Iq and stock market participation. *The Journal of Finance 66*(6), 2121–2164.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review 70*(3), 393–408.

Guiso, L., T. Jappelli, and D. Terlizzese (1996). Income risk, borrowing constraints, and portfolio choice. *The American Economic Review*, 158–172.

Guvenen, F. (2007). Learning your earning: Are labor income shocks really very persistent? *The American economic review 97*(3), 687–712.

Haliassos, M. and C. Bertaut (1995). Why do so few hold stocks? *The Economic Journal 105*(432), 1110–1129.

Haliassos, M. and A. Michaelides (2003). Portfolio choice and liquidity constraints. *International Economic Review 44*(1), 143–177.

Hansen, L. P. and K. J. Singleton (1983). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of political economy 91*(2), 249–265.

Heaton, J. and D. Lucas (2000). Portfolio choice and asset prices: The importance of entrepreneurial risk. *Journal of Finance 55*(3), 1163–1198.

Heaton, J. and D. J. Lucas (1996). Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of political Economy 104*(3), 443–487.

Hong, H., J. Kubik, and J. Stein (2004). Social Interaction and Stock-Market Participation. *Journal of Finance LIX*(1), 137–163.

Hurst, E., M. C. Luoh, and F. P. Stafford (1998). nthe wealth dynamics of american families. *oBrookings Papers on Economic Activity*, 267.

Liu, H. and M. Loewenstein (2002). Optimal portfolio selection with transaction costs and finite horizons. *The Review of Financial Studies 15*(3), 805–835.

Longstaff, F. A. (2001). Optimal portfolio choice and the valuation of illiquid securities. *The Review of Financial Studies 14*(2), 407–431.

Luttmer, E. G. J. (1999). What Level of Fixed Costs Can Reconcile Consumption and Stock Returns? *Journal of Political Economy 107*(5), 969–997.

Malloy, C. J., T. J. Moskowitz, and A. Vissing-Jørgensen (2009). Long-run stockholder consumption risk and asset returns. *The Journal of Finance 64*(6), 2427–2479.

Mankiw, G. and S. P. Zeldes (1991). The consumption of stockholders and nonstockholders. *Journal of Financial Economics 29*, 97–112.

Newey, W. K. and K. D. West (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *Journal of Finance 53*(6), 1887–1934.

Paiella, M. (2001). Limited financial market participation: a transaction cost-based explanation . *Institute for Fiscal Studies: UK IFS Working Paper*.

Paiella, M. (2004). Heterogeneity in financial market participation: appraising its implications for the c-capm. *Review of Finance 8*(3), 445–480.

Paiella, M. (2007). The forgone gains of incomplete portfolios. *The Review of Financial Studies 20*(5), 1623–1646.

Runkle, D. E. (1991). Liquidity constraints and the permanent-income hypothesis: Evidence from panel data. *Journal of Monetary Economics 27*(1), 73–98.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters 20*(2), 177–181.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *The Journal of Finance 56*(2), 433–470.

Vissing-Jørgensen, A. (2002a). Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy 110*, 825–853.

Vissing-Jørgensen, A. (2002b). Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures. *Working Paper, Northwestern University, Kellogg School of Management*.

Vissing-Jorgensen, A. (2004). Perspectives on behavioral finance: Does "irrationality" disappear with wealth? Evidence from expectations and actions. In *NBER Macroeconomics Annual 2003, Volume 18*, pp. 139–208. The MIT Press.

Vissing-Jørgensen, A. and O. Attanasio (2003). Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion. *American Economic Review 93*(2), 383–391.

Wermers, R. (2000). Mutual Fund Perforance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses. *Journal of Finance 55*(4), 1655–1695.

Figure 1: Consumption risk by stockholder classification

This figure depicts the consumption risk for different classifications of stockholders. We use simulated consumption data from the baseline estimation and compute a time-series of the average consumption growth in each period for each stockholder group. We define consumption risk as the standard deviation of these time-series.

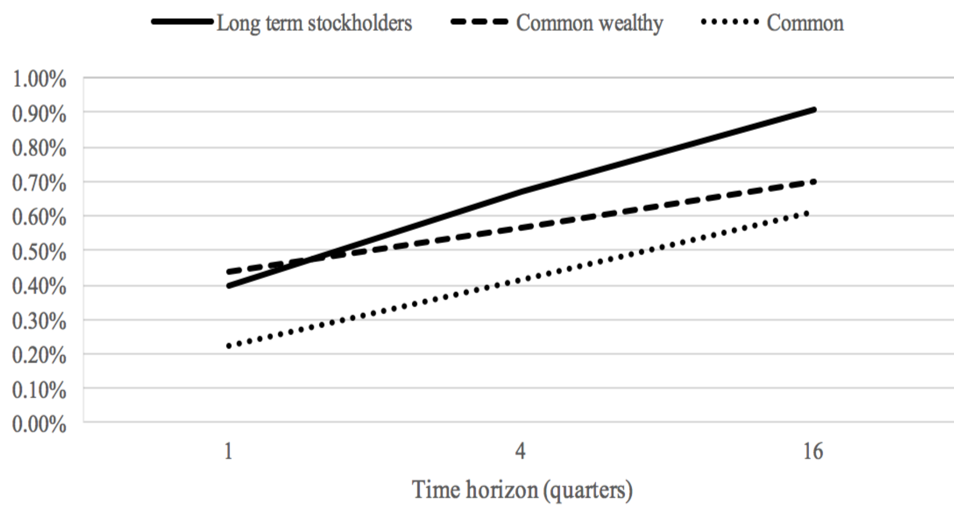

Table 1: Panel Study of Income Dynamics (PSID): 1999-2011

This table reports statistics using the PSID data waves: 1999, 2001, 2003, 2005, 2007, 2009 and 2011. Panel A reports entry and exit statistics. Columns 2 reports the fraction of new stockholders in year $t+2$. That is, the fraction of households that were not stockholders in year t but became stockholders in the following year $t+2$ as a fraction of all households. Columns 2 reports the fraction of households that own stocks in year t but became new non-stockholders in the year $t+2$. Column 3 reports the net new entry, which is column 1 minus column 2. Column 4 reports the fraction of non-stockholders in year t that enter the market in year $t+2$ as fraction of all the non-stockholders in year t. Column 5 reports the fraction of stockholders in year t that exit the market in year $t+2$ as fraction of all the stockholders in year t. In 2001, about 23.9% of stockholders in 1999 exit the market. Panel B reports statistics about the baseline cohort of stockholders from the 1999 wave. Column 1 reports the fraction of this baseline cohort who participate in the stock market in any of the future waves. Column 2 reports the fraction of 1999 stockholder who exit the stock market in a particular wave. Column 3 reports the fraction of the baseline stockholders in year t who participated in the stock market in each of the previous waves. Column 4 reports the fraction of baseline stockholders who re-enter the stock market after having exited in a previous wave. The bottom row in the table reports the averages of each column. Panel C reports the portfolio rebalancing rates for stockholders, where the first row labeled "Year" reports the survey year. The second row reports the average of the trading indicator across households by year.

Panel A: Stock market entry and exit of PSID 1999 cohort

Wave	New Participants (1)	New Non-Participants (2)	Net (3)	Change of status between waves Entry (4)	Exit (5)
1999	-	-	-	-	-
2001	9.7%	7.5%	2.2%	30.9%	23.9%
2003	7.2%	9.9%	-2.7%	21.5%	29.6%
2005	7.2%	8.3%	-1.1%	23.4%	27.0%
2007	7.2%	7.2%	0.0%	24.3%	24.3%
2009	6.3%	8.6%	-2.2%	21.3%	28.9%
2011	5.9%	10.8%	-4.9%	21.6%	39.3%
Average	7.3%	8.7%	-1.5%	23.8%	28.8%

Table 1: Panel Study of Income Dynamics (PSID): 1999-2011 – Cont'd

Panel B: Tracking the stockholders from wave 1999

Wave	Participation of 1999 stockholders in future waves (1)	Exit of 1999 stockholders (2)	1999 stockholders who own stocks in previous waves (3)	1999 stockholders who exited and re-entered (4)
1999	100.0%	-	100.0%	-
2001	76.1%	-23.9%	76.1%	-
2003	69.2%	-6.9%	59.8%	39.2%
2005	65.6%	-3.6%	50.5%	18.6%
2007	64.4%	-1.2%	45.5%	19.3%
2009	58.6%	-5.8%	40.3%	14.8%
2011	50.8%	-7.8%	32.8%	4.5%
Average	69.2%	-8.2%	57.9%	19.3%

Panel C: Trading of stocks

Year	1999	2001	2003	2005	2007	2009	2011	Average
Annual Trading	-	0.561	0.480	0.510	0.476	0.429	0.462	0.486

Table 2: Survey of Consumer Finance: 1989-2010

This table reports estimates of some moments using the SCF data about stock market participation, wealth to income ratio, as well as stockholding to wealth ratio. The SCF waves are 1995, 1998, 2001, 2004, 2007 and 2010. For panel data, we use only the most recent SCF panel for years 2007 and 2009 (we do not use prior SCF panels of 1983-1989 and 1962-1963 because they are outside of our PSID data period). Panel A reports the statistics for entry and exit using the SCF 2007-2009 panel. For the year 2007, below the first column labeled "2007", the rows report the stockholding status of the household as not owning (Not) or owning stocks (Own). For the year 2009, the columns labeled "Not" and "Own", report the stock holdings status for the household. Panel B reports household stock market participation rate, where the first row labeled "Year" reports the survey year. The second row labeled "SCF" reports the stock market participation rate for a given year. Panel C reports summary statistics of wage and equity (stockholdings) for years 2007 and 2009. Specifically, for each sample examined, the columns report the mean and median of equity and wage in 2007 and 2009. The samples include "All households"; those who own equity in the year 2009, "Stockholders 09"; those who own equity in the year 2007, "Stockholders 07"; those who enter the stock market in the year 2009 but were not stockholders in 2007, "Entering Stockholders"; and finally, those who are not stockholders (equity=0) in the year 2009, but were stockholders in 2007 (equity>0), "Exiting Stockholders". Panel D focuses on households who exit the stock market (stockholders in 2007 and not 2009) and those who enter the stock market (stockholders in 2009 but not 2007). The column labeled "Averages" reports the average stockholdings for exiting or entering households. The column labeled "% of 09 equity" reports the average stockholding as a percentage of average equity held by all stockholders in 2009. The column labeled "% of average 09 earnings" reports average stockholdings as a percentage of average labor earnings of stockholders in 2009. Panel E reports ratios for stock holdings, wealth, and wage, where the top row labeled "Survey year" reports the survey year. Panel E.1 reports statistics for stockholder households, where the first row labeled "Equity share" reports the median stock holdings divided by financial wealth. The second row labeled "Median of Fin. Wealth to Wages (Stockholders)" reports the median of financial wealth divided by wage. Panel E.2 reports statistics for all households (stockholder and non-stockholder), where the first row labeled "Median of Fin. Wealth to Wages (All households)" reports the median wealth to wage ratio for all households.

Panel A: Entry and exit- SCF 2007-2009

Year/Year:	2009	
2007	Not	Own
Not	37.1%	9.3%
Own	7.4%	46.2%

Panel B: Stock market participation SCF 1989-2010

Year	1995	1998	2001	2004	2007	2010	Average
SCF	0.411	0.489	0.522	0.502	0.511	0.499	0.489

Table 2: Survey of Consumer Finance: 1989-2010 – Cont'd

Panel C: Summary statistics of wage and equity

	All households		Stockholders 09		Stockholders 07		Entering Stockholders		Exiting Stockholders	
Statistics	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Equity 07	124,287	932	217,048	25,890	231,641	36,246	.	.	51,824	9,320
Wage 07	55,512	34,079	79,765	56,444	81,202	57,509	43,888	34,079	45,425	30,884
Equity 09	87,771	1,500	158,189	27,500	159,721	25,000	22,350	5,000	.	.
Wage 09	53,932	32,875	77,846	58,776	77,638	57,780	47,856	37,856	38,991	25,901

Panel D: Exit and entry volume

Status	Average ($)	% of average 09 equity	% of average 09 earnings
Exit	51,824	32.8%	66.6%
Entry	22,350	14.1%	28.7%

Panel E: Financial wealth and wage: stockholders (panel E.1) and all households (panel E.2)

Survey year	1989	1992	1995	1998	2001	2004	2007	2010	Average
	Panel E.1: Share of stocks from financial assets (stockholders)								
Stock/Fin	0.27	0.32	0.41	0.48	0.51	0.46	0.44	0.40	0.41
Fin/Wage	0.79	0.88	0.86	1.06	1.14	1.20	1.20	1.14	1.03
	Panel E.2: Median financial assets by wage (all households)								
Fin/Wage	0.29	0.30	0.37	0.50	0.53	0.47	0.54	0.40	0.42

Table 3: Probit estimates and persistence in participation

This table reports summary statistics (Panel A) and estimates from probit regressions (Panel B). The dependent variable in the probit regressions is a binary variable for current stock market participation (1 if the household owns stocks, 0 otherwise). The key independent variable is a binary variable for the previous stock market participation status and labeled as "Past participation". The other independent variables are household demographics such as age, gender and race. Panel A reports the number of observations, mean, standard deviation, and median of the key variables from both the SCF and PSID used in the probit regressions. In Panel B, regression 1 reports results using the SCF panel 2007-2009. Regressions 2 to 4 use the PSID data, where regression 2 uses only the PSID panel 2007-2009, and regressions 3 and 4 use the entire sample 1999-2011. Regression 3 (4) omits (includes) year fixed effects. The numbers in parentheses are the z-statistics based on robust standard errors.

Panel A: Summary Statistics

	SCF				PSID			
	Households	mean	sd	median	Obs	mean	sd	median
Past Participation	3,857	0.537	0.499	1	203,586	0.221	0.415	0
White	3,857	0.736	0.441	1	178,206	0.772	0.419	1
Male	3,857	0.719	0.45	1	178,206	0.774	0.418	1
Age	3,857	5.15	16.7	50	203,586	45.2	18.6	43
Education	3,857	13.371	2.769	13	175,259	13.388	2.767	13
Income ($mn)	3,857	0.056	0.127	0.034	177,002	0.08	0.097	0.06
Wealth ($mn)	3,857	0.209	1.486	0.024	165,991	0.277	0.982	0.051

Table 3: Probit Estimates and persistence in participation – Cont'd

Panel B: Probit Regressions				
	SCF	PSID	PSID	PSID
	(1)	(2)	(3)	(4)
Past Participation	0.575	0.507	0.483	0.481
	(97.35)	(91.54)	(227.20)	(226.14)
White	0.056	0.039	0.047	0.047
	(9.05)	(7.14)	(21.71)	(21.78)
Male	0.062	0.007	0.025	0.025
	(10.25)	(1.32)	(11.69)	(11.91)
Age (/100)	-0.046	0.190	0.148	0.154
	(-2.76)	(14.55)	(30.32)	(31.62)
Education	0.026	0.015	0.017	0.017
	(24.62)	(17.78)	(50.74)	(51.19)
Income (mn)	0.159	0.244	0.218	0.220
	(7.08)	(9.85)	(21.34)	(21.52)
Wealth (mn)	0.003	0.000	0.000	0.000
	(1.85)	(10.88)	(31.50)	(31.55)
Constant	-0.170	-0.261	-0.270	-0.287
	(-9.85)	(-18.69)	(-50.14)	(-49.52)
Fixed effect	-	-	No	Yes
Households/[Obs]	3,857	[22,568]	[159,082]	[159,082]
Pseudo R^2	0.470	0.394	0.362	0.365

Table 4: Cross-sectional determinants of entries and exits

This table reports estimates from cross-sectional regressions. Panel A shows summary statistics of the variables in these regressions and Panel B reports the regression results. The dependent variables in regressions 1, 2 and 3 are the log of 1 plus the total number of entries, exits, and entries and/or exists, respectively. In regressions 4, 5, and 6, the dependent variables are dummy variables that take the value of 1 if over the sample period the household had at least one entry, one exit, and one entry or exit, respectively. We estimate regressions 1 to 3 with ordinary least squares and report the respective estimates and t-statistics. The latter are reported in parenthesis underneath the coefficient estimates and are based on robust standard errors. We estimate regressions 4 to 6 with a probit estimator and report the respective marginal effects. The z-statistics are also based on robust standard errors. For the estimation, we use data from the 1999 to 2011 PSID waves. We exclude from the sample households that never participated in the stock market in the 1999 to 2011 period and households with income growth higher than 300% and lower than -70%. The independent variables are average income, average wealth (i.e., net worth), and income risk (computed as the standard deviation of biennial income growth between 1999 to 2011). The control variables are age, race, gender, and education as of 2011. The definition of all variables are in Appendix D.

Panel A: Summary Statistics

			Percentile					
variable	mean	sd	p10	p25	p50	p75	p90	N
Log of (1 + entries)	0.522	0.399	0	0	0.693	0.693	1.099	4,085
Log of (1 + exits)	0.605	0.389	0	0	0.693	0.693	1.099	4,085
Log of (1 + exits/exits)	0.907	0.51	0	0.693	1.099	1.099	1.609	4,085
Entries	0.819	0.693	0	0	1	1	2	4,085
Exits	0.964	0.702	0	0	1	1	2	4,085
Entries and Exits	1.783	1.254	0	1	2	2	4	4,085
I(Entries>0	0.665	0.472	0	0	1	1	1	4,085
I(Exits>0)	0.747	0.435	0	0	1	1	1	4,085
I(Entries/Exits>0)	0.821	0.383	0	1	1	1	1	4,085
Average income (mn)	0.742	0.667	0.275	0.413	0.622	0.882	1.255	4,085
Average wealth (mn)	0.534	1.214	0.049	0.121	0.275	0.573	1.11	3,984
Income risk (/100)	0.005	0.008	0.001	0.002	0.003	0.005	0.008	4,085
Age	57.09	13.896	40	46	56	66	78	4,085
White	0.907	0.29	1	1	1	1	1	4,085
Male	0.893	0.309	0	1	1	1	1	4,085
Education	14.597	2.229	12	12	16	16	17	4,055

Table 4: Cross-sectional determinants of entries and exits – Cont'd

Panel B: Cross-sectional Regressions						
	(1)	(2)	(3)	(4)	(5)	(6)
	Log(1 + Entries)	Log(1 + Exits)	Log(1 + Entries/Exits)	I(Entries>0)	I(Exits>0)	I(Entries/Exits>0)
Average income (mn)	-0.013	-0.026	-0.034	-0.007	-0.043	-0.034
	(-0.91)	(-1.61)	(-1.48)	(-0.29)	(-2.39)	(-2.33)
Average wealth (mn)	-0.041	-0.025	-0.048	-0.093	-0.013	-0.016
	(-4.56)	(-1.86)	(-2.68)	(-4.33)	(-1.13)	(-1.56)
Income risk (/100)	2.638	1.062	2.695	9.477	1.546	5.206
	(3.72)	(1.80)	(3.46)	(4.69)	(1.43)	(3.16)
Age	-0.003	-0.002	-0.004	-0.004	-0.004	-0.005
	(-5.95)	(-3.11)	(-5.70)	(-5.91)	(-5.38)	(-8.74)
White	-0.067	-0.059	-0.087	-0.081	-0.095	-0.054
	(-2.81)	(-2.62)	(-2.83)	(-2.79)	(-3.43)	(-2.37)
Male	-0.025	-0.04	-0.051	-0.018	-0.061	-0.038
	(-0.90)	(-1.59)	(-1.49)	(-0.53)	(-1.94)	(-1.52)
Education	-0.027	-0.032	-0.046	-0.037	-0.046	-0.045
	(-8.33)	(-9.88)	(-10.82)	(-8.31)	(-10.89)	(-11.82)
Constant	1.204	1.286	1.979			
	(19.15)	(20.43)	(23.89)			
Observations	3,954	3,954	3,954	3,954	3,954	3,954
R2/[Pseudo R2]	0.066	0.058	0.086	[0.074]	[0.074]	[0.125]

Table 5: Simulated method of moments estimates: Baseline estimation

This table reports the estimation results of the model by SMM. Panel A reports the estimated parameters, where column 1 labeled "Parameter" reports the parameter notation. Column 2 labeled "Coefficient" reports the coefficient estimates. Column 3 labeled "t-statistics" reports the t-statistics. The last row of the table reports the fit of the model based on the J-statistic. Panel B describes the moments of the model, where column 1 labeled "Moment" reports the name of the moments. Column 2 labeled "Data" reports moment values calculated from the data. Column 3 labeled "Baseline estimation" reports the values of model simulation moments.

Panel A: Estimated parameters		
Parameter	Coefficient	t-statistic
γ	3.176	3.41
β	0.963	5.02
Ψ	0.968	4.55
F^A	0.020	6.23
R (annualized)	0.084	5.45
J-statistic	0.104	
P-value of J test	0.010	
Panel B: Moments		
Moment	Data	Baseline estimation
Probit Est(Current Part | Past Part)	0.510	0.443
Entry/Exit	0.160	0.286
Always Participate	0.328	0.146
Participated in first and last year	0.508	0.607
Average Participation in a year	0.492	0.613
Rebalance rate (average of trading indicator)	0.486	0.599
α_1- 1 quarter	0.029	0.029
α_1- 4 quarters	0.127	0.050
α_1- 16 quarters	0.352	0.115
Equity share	0.413	0.784
Median of Fin. Wealth to Wages (Stockholders)	1.034	0.886
Median of Fin. Wealth to Wages (All households)	0.425	0.517
Risk premium (annualized)	0.076	0.074

Table 6: Sensitivity to transaction costs

This table reports how the moments that we match in the SMM estimation respond to variations in the transaction cost parameters. In column 1 and 2 we report the data moments and implied moments from the baseline estimation in Table 5, respectively. In column 3 we report the moments when there are no transaction costs (i.e., $F^A = 0$; $C = 0$; $\Psi = 1$). In column 4 we allow only for fixed participation costs (i.e., $F^A \neq 0$; $C = 0$; $\Psi = 1$). In column 5 we allow for fixed participation costs and the opportunity time cost of rebalancing (i.e., $F^A \neq 0$; $C = 0$; $\Psi \neq 0$). In column 6 we allow for fixed participation costs and monetary trading costs (i.e., $F^A \neq 0$; $C \neq 0$; $\Psi = 1$).

	(1)	(2)	(3)	(4)	(5)	(6)
	Data	Baseline	$F^A = 0$; $C = 0$ $\Psi = 1$	$F^A \neq 0$; $C = 0$ $\Psi = 1$	$F^A \neq 0$; $C = 0$; $\Psi \neq 0$	$F^A \neq 0$; $C \neq 0$; $\Psi = 1$
Probit Est(Current Part | Past Part)	0.510	0.443	0.061	0.061	0.429	0.078
Entry/Exit	0.160	0.286	0.142	0.142	0.291	0.186
Always Participate	0.328	0.146	0.547	0.547	0.160	0.459
Participated in first and last year	0.508	0.607	0.966	0.966	0.633	0.950
Average Participation in a year	0.492	0.613	0.962	0.962	0.630	0.946
Rebalance of Equity Share	0.486	0.599	1.000	1.000	0.622	0.999
α_1- 1 quarter	0.029	0.029	0.035	0.035	0.033	0.034
α_1- 4 quarters	0.127	0.050	0.079	0.079	0.058	0.078
α_1- 16 quarters	0.352	0.115	0.068	0.068	0.078	0.052
Equity share	0.413	0.784	1.000	1.000	0.790	1.000
Median of Fin. Wealth to Wages (Stockholders)	1.034	0.886	0.711	0.711	0.874	0.721
Median of Fin. Wealth to Wages (All households)	0.425	0.517	0.595	0.595	0.522	0.596
Risk premium	0.076	0.074	0.074	0.074	0.074	0.074

Table 7: Cross-sectional household regressions for participation, entry and exit

This table reports regression results at the household level. The dependent variable in regressions (1), (2) and (3) are the fraction of periods the household enters, exit, and either enters or exit the stock market. The dependent variable in regressions (4) and (5) are the fraction of periods the household participates in the stock market. The first independent variable is household income risk, measured as the standard deviation of income growth, labeled "Income risk (st. dev Inc gt)". The second independent variable is the average income level labeled "Income." The numbers in parentheses measure the t-statistics based on robust standard errors.

	(1) Entry	(2) Exit	(3) Entry/Exit	(4) Participation	(5) Participation
Income risk	0.070	0.062	0.131		0.504
	(2.63)	(2.31)	(2.48)		(1.97)
Income	-0.089	-0.094	-0.184	5.927	5.930
	(-6.33)	(-6.63)	(-6.50)	(34.19)	(34.21)
Constant	0.096	0.102	0.198	-5.317	-5.415
	(6.55)	(6.88)	(6.75)	(-30.67)	(-28.23)
Observations	500	500	500	500	500
R-squared	0.101	0.107	0.105	0.690	0.691

Table 8: Reaction of households to exogenous shocks

This table reports estimated simulated moments with an exogenous shock to income risk. Column 2 labeled "Baseline" reports the baseline model moments. Column 3 labeled "High Risk" shows estimated moments when household income risk is increased by 25% compared to the baseline case.

Moment	Baseline Estimation	High risk
Probit Est(Current Part | Past Part)	0.443	0.403
Entry/Exit	0.286	0.263
Always Participate	0.146	0.260
Participated in first and last year	0.607	0.748
Average Participation in a year	0.613	0.730
Rebalance rate (average of trading indicator)	0.599	0.719
α_1- 1 quarter	0.029	0.039
α_1- 4 quarters	0.050	0.099
α_1- 16 quarters	0.115	0.142
Equity share	0.784	0.814
Median of Fin. Wealth to Wages (Stockholders)	0.886	1.239
Median of Fin. Wealth to Wages (All households)	0.517	0.841
Risk premium	0.074	0.074

Table 9: Aggregate time-series results for stock market participation, entry and exit

This table reports time series regressions, where the dependent variable in regression (1) labeled "Avg Entry" measures the fraction of households who enter the stock market. The dependent variable in regression (2) labeled "Avg Exit" measures the fraction of households who exit the stock market. The dependent variable in regression (3) labeled "Avg Participation" measures the fraction of households who participate in the stock market. And, the dependent variable in regression (4) labeled "Avg Flow" measures the net cash flow entering the stock market. The independent variables are the stock market return R_t, and the average income growth g_t. The numbers in parentheses measure the t-statistics based on robust standard errors.

	(1) Avg Entry	(2) Avg Exit	(3) Avg Participation	(4) Total Flow
R_t	-0.034	-0.030	0.396	166.940
	(-3.16)	(-2.92)	(4.60)	(40.45)
Income g_t	0.078	-0.071	-0.065	52.394
	(2.61)	(-2.49)	(-0.27)	(4.52)
Constant	0.020	0.020	0.604	12.019
	(62.56)	(62.61)	(227.03)	(94.33)
No. of time periods	799	799	799	799
R-squared	0.020	0.019	0.026	0.677

Table 10: Different stockholder classification and estimation of the EIS

This table reports financial summary statistics and estimates of EIS by stockholder classification (long term stockholders, common wealthy, and common). All the data are from the simulation of our model in Section 3 under the baseline parameter estimates in Table 5. The "common" stockholders are those that participate in a given period regardless of their participation status in other periods. The "common wealthy" are the top one third wealthiest common stockholders. The "long-term stockholders" are the common stockholders that remain in the stock market for at least 70% of periods in our sample. The financial summary statistics include financial wealth and income relative to their averages, income risk measured as the standard deviation of income growth, and the rebalancing rate. To compute the rebalancing rate, we first compute the time series average of the trading index of each household and then average across all households in each stockholder group. The last row reports the estimated EIS and its respective t-statistic between parenthesis.

Stockholder classification	Long term stockholders	Common wealthy	Common
Relative financial wealth	1.312	1.541	1.239
Relative income	1.030	1.042	1.019
Income risk	0.023	0.032	0.012
Rebalancing rate (avg. of trading indicator)	0.717	0.652	0.323
Estimated EIS	0.236 (2.87)	0.176 (2.61)	0.111 (2.04)

Appendix

A Overview of stockholder classifications

In this appendix we present a summary of how the literature has dealt with defining who the stockholders are. From our reading of the literature, we concluded that this definition is highly variable across studies. It depends on the goal of study as well as the data being used. Therefore, some researchers distinguish between stockholders with non-retirement holdings (e.g. holding shares of publicly held corporations or mutual funds) and retirement holdings (e.g. pension plans and IRA's), while others do not.

In principle, it is difficult to justify treating retirement and non-retirement holdings equivalently. As Haliassos and Bertaut (1995) note, "*membership and dependence of pension income on stock market performance constitute a state-contingent claim with different liquidity properties and payoffs than direct stock ownership.*"

Despite the greater autonomy of the account holder, IRAs are consistently treated like pension plan accounts (Ameriks and Zeldes 2004b). This is the case since the accounts include restrictions similar to those of pension accounts. In particular, holders are given incentives to contribute to the account and receive tax benefits when the contributions remain in the account. Moreover, similar to 401(k) accounts, IRAs' distributions begin at age 59.5, are forced when the account holder does not withdraw anything at age 70.5, and are subject to penalties if withdrawals come early. It follows that, if pension accounts are excluded from the definition of stock holding because of their distinctive properties, then IRA's should be treated in a similar manner.

In many studies the definition of stockholders depends on the source of the data used, and whether there is information to distinguish between different types of holdings. Studies that employ the Consumer Expenditure Survey (CEX) are not able to distinguish between retirement and non-retirement holdings because the survey question related to ownership of risky assets does not separate the two (Vissing-Jørgensen 2002b; Vissing-Jørgensen and Attanasio 2003; Paiella 2001; Brav, Constantinides, and Geczy 2002).

In the PSID, one can identify whether a household owns risky assets in non-retirement accounts in the survey waves administered from 1999 and onward. This information can be extracted from a new question added in 1999. Before 1999, it is not possible to separate direct stock ownership from stock ownership via retirement accounts (e.g., Mankiw and Zeldes (1991)). Moreover, prior to 1999, information on combined market participation (i.e., direct and/or indirect) is only available in the 1984, 1989, and 1994 waves. Even after 1999, the PSID does not provide wealth information to separate the portion of wealth allocated to retirement and non-retirement investment accounts.

In the case of the Health and Retirement Study (HRS), one can separate between non-retirement and retirement holdings (Haliassos and Bertaut 1995; Hong, Kubik, and Stein 2004; Bogan 2008). However, the HRS has little information on assets held in retirement accounts. Therefore, as noted by Hong, Kubik, and Stein (2004), stockholders can be defined by a survey question on whether households own stocks, either directly or through mutual funds, which only pertains to non-retirement investment accounts.

The Survey of Consumer Finances (SCF) provides the most detailed information of all the surveys. In particular, studies using the SCF have used both definitions of stockholders, as those with both non-retirement and retirement holdings (Haliassos and Bertaut 1995; Ameriks and Zeldes 2004b), as well as those with only non-retirement holdings (Bogan 2008). Moreover, the SCF provides information on wealth and one can separate the portion of wealth in retirement and non-retirement investment accounts.

B Estimation of proportional trading costs function

In this appendix, we provide the estimates of the proportional trading cost function. The estimation follows Bonaparte, Cooper, and Zhu (2012). As in Bonaparte, Cooper, and Zhu (2012), we assume that overall trading costs are a quadratic function of the trade value v.

Table B1: Estimation trading costs function

This table reports regression results for the cost of trading (buying and selling stock), where the dependent variable is the commission, and the independent variables are trade value (the price of the share times the quantity of share) and trade value squared. If a stockholder buys different stocks in a given month, the stockholder reports the commission, quantity and price on each one of these stocks separately. The data is from the Barber and Odean (2000) study that contains information on common stock trades of about 78,000 stockholders who are clients of a discount brokerage firm from January 1991 to December 1996. Finally, the numbers in parentheses are t-statistics, which are based on standard errors clustered at the account level.

Parameter	Buying	Selling
Constant v_0^i	56.106	61.437
	(64.32)	(129.05)
Linear v_1^i	0.001	0.001
	(14.69)	(36.72)
Quadratic v_2^i	-2.88E-13	-9.26E-13
	(-5.78)	(-2.43)
Observations	1,746,403	1,329,394
R-squared	0.251	0.359

C Estimation of household income process

When we estimate the household income process we follow the life-cycle literature (e.g., Viceira (2001), Gourinchas and Parker (2002)) and assume that income is affected by a deterministic component, which depends on demographic characteristics, and by a stochastic component. In particular,

$$y_{i,t} = \tau Z_{i,t} + A_{i,t}.$$

Above, $y_{i,t}$ is labor income for household i at time period t. The term $\tau Z_{i,t}$ is the deterministic component of income. $Z_{i,t}$ is a vector of household demographic variables like age, and τ is the corresponding vector of coefficients. $A_{i,t}$ is the stochastic component of income. A is persistent and follows the process $A_{i,t} = \rho A_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,t}$ is the transitory shock. The above specification is also similar to that in Campbell, Cocco, Gomes, and Maenhout (2001).

We estimate the above income process using data from the Panel Study on Income Dynamics (PSID). Our sample period is from 1967-1993. We use data till 1993 since the PSID has annual surveys till 1993; after that the surveys are administered every two years. In the estimation, we deflate the labor income levels using the CPI obtained from the Bureau of Labor Statistics.

We estimate the income process using a restricted sample where the head of the household is: (i) male,[22] (ii) between 20 and 64 years old, and (iii) not from the SEO sample. We also require that the real hourly labor earnings of the head of the household is between \$2 and \$400. Finally, we focus on households where the head of household works between 520 hours (10 hours per week) and less than 5,110 hours (14 hours a day, every day).

[22] We focus on males following the income profile literature. For instance, see Guvenen (2007)

We estimate the income process following the two step approach of Bonaparte, Cooper, and Zhu (2012). In particular, we first pool the observations across all individuals and regress income on demographics, such as age, age squared and education attainment. We treat the explained part of this regression as the deterministic component of income. Like Carroll and Samwick (1997), we use the error term from the demographic-based regression to capture the unobservable stochastic component of income. As noted above, we assume that the stochastic component of income follows an AR(1) process. We fit the AR(1) process to the residuals and find that the autocorrelation parameter ρ is 0.842. Finally, from the estimation of the AR(1) process, we obtain the standard deviation of the innovation ϵ, which is 0.290.

After the estimation, we use the autocorrelation of the permanent component A and the standard deviation of ϵ to simulate income data for the households in our model. In particular, for the simulation of the model, we transform the estimated income process to a 5-state Markov chain following the Tauchen (1986) methodology. Also, in the Markov chain, we set the average level of real income to $72,000 annually to mimic the average income level in our PSID sample.

D Definitions of key empirical variables

In this appendix, we provide detailed definitions of our variables. In the PSID, we use the waves between 1999 to 2011, which are administered every two years. To mitigate measurement error concerns, we exclude from our PSID households with inconsistent information about age, gender and race. For example, we exclude households that the gender of the head of household is different across waves. In the case of the SCF, we use the panel from the 2007 and 2009 waves. Because the SCF over-samples wealthy households, we follow the SCF guidelines and use the sample weights provided by the SCF to ensure that our estimates (i.e., summary statistics and regression estimates) are representative of the average U.S. household.

Table D1: Variable definitions

The table presents definitions of the variables extracted from the SCF and the PSID

Variable	Description	Source
Stock market participation	1 if own NON-IRA stocks, 0 otherwise, based on the answer to: *Do [you/you or anyone in your family] have any shares of stock in publicly held corporations, mutual funds, or investment trusts – not including stocks in employer based pensions or IRAs?*	PSID
Past participation	1 if own NON-IRA stocks in the past wave, 0 otherwise	PSID
Stock market participation	1 if portion of wealth allocated to NON-IRA stocks>0, 0 otherwise	SCF
Past participation	1 if portion of wealth allocated to NON-IRA stocks>0 in the past wave, 0 otherwise	SCF
Equity share	Stock holding divided by financial assets	SCF
Financial Wealth to Wages	Value of financial assets divided by labor income	SCF
Trading Indicator	1 if the household buy/sell stocks, 0 otherwise (based on the answer to question whether respondent bought or sold any non-IRA stocks since the previous interview)	PSID
Rebalancing rate	Average of the trading indicator	PSID
Entry	1 if the household enters the stock market in wave t, 0 otherwise (i.e., household owns stocks in wave t but not in the previous wave)	PSID/SCF
Exit	1 if the household exits the stock market, 0 otherwise (i.e., household does not own stocks in wave t but owned stocks in the previous wave)	PSID/SCF
Entry/Exit	1 if the household entry or exit dummy variable is 1, 0 otherwise	PSID/SCF

Table D1: Variable definitions – Cont'd

Variable	Description	Source
White	1 if the household race is white, 0 otherwise	PSID/SCF
Male	1 if the household gender is male, 0 otherwise	PSID/SCF
Age	Years old	PSID/SCF
Education	Years schooling	PSID/SCF
Wage/Income	Labor earnings	PSID/SCF
Wealth	Assets - Liabilities	PSID/SCF
Always Participate	Fraction of households who own stocks in all waves	PSID
Participated in first and last year	Fraction of households who owns stock in 1999 and in 2011	PSID
Log of (1 + entries)	The log of 1 plus the total number of entries in the stock market between 1999 and 2011	PSID
Log of (1 + exits)	The log of 1 plus the total number of exists from the stock market between 1999 and 2011	PSID
Log of (1 + exits and exists)	The log of 1 plus the total number of entries and exists from the stock market between 1999 and 2011	PSID
I(Entries>0)	1 if household enters the stock market at least once, 0 otherwise	Created
I(Exits>0)	1 if household exits the stock market at least once, 0 otherwise	Created
I(Entries/Exits>0)	1 if household enters or exits the stock market at least once, 0 otherwise	Created
Average income	The average income between 1999 and 2011	PSID
Average wealth	The average wealth between 1999 and 2011	PSID
Income risk	The standard deviation of the biennial income growth between 1999 and 2011	PSID

E Literature review on transaction costs

In this appendix, we discuss in detail the transaction costs components that we include in our model. We also provide a review of the literature on how these costs are modeled.

In classifying the components of transaction costs in our model, we follow Vissing-Jørgensen (2002b) and categorize them in three groups. They are fixed-type costs of trading, proportional costs of trading, and per-period time-costs to stock ownership. The first cost component is fixed-type costs of trading. For stockholders who trade, this represents the costs of maintaining the trading accounts and any other fixed brokerage fees and expenses.

The second cost component is proportional costs of trading. This cost component is a variable cost. For those who own mutual funds, this would reflect the front load paid on entry into the fund. Vissing-Jorgensen (2004) also suggests that "some funds have contingent deferred sales loads requiring investors to pay a certain percentage of their initial investment if they sell their mutual fund shares before a given number of years." On the other hand, households with direct stock holdings incur this cost through the bid-ask spread and the variable portion of brokerage commissions and fees.

The final cost component is per-period time-cost to stock ownership. For households that directly hold stocks, this cost includes the value of time spent throughout the year determining if trading is optimal. With time varying conditional asset distributions, theory suggests that households should actively follow the stock market to form more precise expectations of future returns and rebalance their portfolios accordingly. Moreover, Vissing-Jorgensen (2004) reports that "according to the Internal Revenue Service (IRS) numbers for 2002, households who have to fill out schedules D and D1 (the schedules for capital gains and losses) spend 8 hours and 34 minutes on average doing so.

For those households who indirectly hold stocks through mutual funds, the per-period time cost represents the cost of delegating the role of portfolio manager to the mutual fund manager. Such delegation costs typically represent the annual expense fees charged to investors. Wermers (2000) estimates average expense ratios of mutual funds at around 0.93% per year for the period of 1990-1994 based on actively managed mutual funds (averaged over total net assets of funds).

Next, we provide a short literature review of the literature that includes transaction costs in portfolio choice models. The literature models transactions costs in various ways. Despite the distinct components described above, most studies employ a lump sum cost that is paid once. For instance, Gomes and Michaelides (2008) and Alan (2006) employ a life cycle model and charge households a one-time cost that is between 2.22% and 2.50% of the permanent component of labor earnings.

Luttmer (1999) and Paiella (2001) emphasize the per-period costs needed to prevent households from adjusting their consumption from its current value in order to participate in the market. Luttmer (1999) proposes a fixed transaction cost at a minimum of 3% of monthly per capita consumption, and suggests a variable cost of roughly $15. Paiella (2001) suggests that a yearly cost of at least $31 is needed to rationalize nonparticipation for a consumer with log utility who can trade US Treasury Bills. Similarly, Paiella (2007) finds that the lower bound of fixed costs that can rationalize not participating in the stock market is $130 per year.

Attanasio and Paiella (2011) generalize the model in Paiella (2007) and estimate preference parameters and a lower bound for market participation costs. Their goal is to jointly reconcile the equity premium puzzle and limited stock-market participation. The use data from the Consumer Expenditure Study (CEX) and estimate their model with the generalized method of Moments (GMM). Their GMM estimate of the coefficient of relative risk aversion is 1.7 and the implied lower bound for participation costs is about $72 per year.

This cost represents about 0.4% of non-durable consumption.

Vissing-Jorgensen (2004) uses Euler equations related to participation and non-participation to indirectly infer total participation costs. She finds that overall costs are as high as $800 for the median household per year (in 1984 dollars). She also shows that the estimated cost varies by household, and depends on the fraction of risky assets in the portfolio.

Finally, in terms of delegated investments such as mutual funds, Wermers (2000) decomposes the associated transaction costs for funds into the trade-related costs incurred by the fund, and the expense ratio charged by the fund. The author finds that over the 1975 - 1994 period, trading costs (per dollar invested in mutual funds) in 1994 are about one-third of their level in 1975. This is ascribed to declining market costs to trade partly driven by technological advancements. However, the average expense ratio over the same period has increased (as a percentage of assets). This increase is explained by a larger proportion of new small funds in later periods, as well as a substitution of 12b-1 fees for load fees during the 1990s.

F Definitions of key theoretical parameters and functions

In this appendix, we provide detailed definitions of the variables and functional forms used in the model in section 2.

Table F1: Key theoretical parameters and functions

The table provides the definitions of variables and function form the model in Section 2.

Parameter/Function	Description
E	Expectation operator
$v(.)$	Value function
$u(.)$	Utility function - CRRA
γ	Coefficient of relative risk aversion
β	Discount factor
con	Non-durable consumption
s	Risky asset (stocks) holdings
b	Riskless asset (bond) holdings
y	Labor earnings
$1-\Psi$	Portion of current income lost due to the time spent rebalancing the portfolio
F^A	Fixed-type trading costs (e.g., account maintenance fees)
$C(.)$	Costs of trading, including commissions, fees, and bid-ask spread
r	Risk free return
R	Risky stock market return
α_1	Sensitivity of aggregate consumption growth to stock returns ("aggregate" EIS)

G Model computation

We implement the simulated method of moments estimation by solving the model with a value function iteration approach. Below, we provide the details of this methodology.

To begin with, the state space of the dynamic optimization problem is determined by y, s_{-1}, b_{-1}, and R_{-1}. y is income of the current period; s_{-1} and b_{-1} are the beginning-of-period asset holdings of risky stocks with return R_{-1} and riskless bonds with return r, respectively. In solving he model, we make the following assumption about the income and return processes. First, as mentioned in Appendix C, we transform the estimated income process to a 5-state Markov chain following the Tauchen (1986) methodology. Second, we assume that stock return process is an IID process with two return states. Its quarterly standard deviation of 8.3%. For simplicity, we require the stock return and the income process to be uncorrelated. Third, we fix the return of the risky-free asset to be 1% per annum (or 1.0%/4=0.25% on a quarterly basis).

In the model, at the beginning of each period, a stockholder makes the decision of how much of his wealth and income to consume and how much to allocate to stocks and bonds. This is a high-dimensional programming problem and it is computationally intensive. Thus, to solve the model with good precision within a reasonable amount of time, we implement the following strategy to solve for the decisions of a stockholder:

1. We assume that the choice of stock holdings (control variable) is made before the realization of return of the risky asset. Thus, in our model simulation, we first solve for the control variable of stock holdings and then multiply it with the return to make it a state variable for the next period.

2. We utilize a mixture of grid search and spline interpolation to execute the value function iteration. In particular, we define a coarse grid with 25 points for stock

holdings and 20 points for bond holdings, denoted $s_{coarse} \times b_{coarse}$. Then, we turn to a fine grid with 400 150 grid points, denoted as $s_{fine} \times b_{fine}$.

Finally, to operationalize the value-function iteration, we guess the value function values for each discrete state variable, and then update the value function as follows:

1. We compute the value of the sub-optimal decision of not adjusting stock holdings, denoted by $v^n = (D, s_{coarse} \times b_{coarse})$, where D stands for the product of discrete state variables on the coarse grid. The value of the sub-optimal decision of always adjusting stock holdings, denoted by $v^\alpha = (D, s_{coarse} \times b_{coarse})$, is also computed.

2. We then use the values of $v^n = (D, s_{coarse} \times b_{coarse})$ and $v^\alpha = (D, s_{coarse} \times b_{coarse})$ to interpolate the values on the fine grid, denoted $v^n = (D, s_{fine} \times b_{fine})$ and $v^\alpha = (D, s_{coarse} \times b_{coarse})$. For the interpolation, we use spline interpolation since the value function is highly non-linear.

3. Finally, we compute the updated value function as:

$$v(D, s_{fine} \times b_{fine}) = max\{v^n(D, s_{fine} \times b_{fine}), v^\alpha(D, s_{fine} \times b_{fine})\}.$$

The policy function and the simulated data are computed on the fine grid after the convergence of the value function. The coarse grid and fine grid are designed cautiously since high upper bounds reduce the efficiency of the optimization routine while low upper bounds cause stockholders optimal decision rules to be distorted. To address these issues, and based on several experiments, we placed more points near the lower-bounds of the asset holdings grid. Specifically, the optimal upper-bound for stockholdings is 40 times the mean income, and 20 times the mean income for bond holdings. We imposed a lower bound on assets to be zero, so there is no shorting. Finally, we draw random shocks to income and returns for 500 stockholders for 800 periods (quarters), so the dimension of simulated data is 500 $\times$ 800.

H Sensitivity analysis

In this appendix, we present the sensitivity analysis related to the SMM estimation from Section 4.5. We report our results in Table H1 where we show how the model-implied moments respond to variations in preference parameters. This analysis reveals whether the model inference is sensitive to small changes in the deep parameters. Specifically, in Table H1, we report the elasticity of the moments (columns) to variations in the parameters (rows). We use the numerical derivatives computed in the neighborhood of the baseline parameter estimates and compute the percentage change of a moment with respect to 1% percentage change in a parameter. For comparison, we also report in column 2 the moments from the baseline estimation.

The results in Table H1 show that overall the model-implied moments are not influenced by changes in the deep parameters of the model. The only exception is the case of the discount factor β. In particular, changes in the discount factor β strongly affect the aggregate EIS α_1, and the entry/exit moments. The discount factor β also affects the wealth to income ratio. The higher the discount factor, the higher the accumulation of wealth relative to income. Furthermore, we find that the rebalancing rate is very sensitive to the rebalancing cost. Finally, most moments are not sensitive to changes in risk aversion γ, except the return elasticity moments where the change is large for α_1 in the 16 quarter time horizon.

Table H1: Elasticities of moments to parameters

This table reports how the moments respond to variations in parameters. These numerical derivatives are computed in the neighborhood of the baseline parameter estimates. The five parameters are listed as columns, and the twelve moment's row. An entry is an elasticity: the percentage change of a moment with respect to the one percentage change in a parameter.

Moment	Baseline Estimation	γ	β	Ψ	F^A	R
Probit Est(Current Part | Past Part)	0.443	0.447	0.453	0.443	0.443	0.447
Entry / Exit	0.286	0.282	0.181	0.286	0.286	0.285
Always Participate	0.146	0.148	0.452	0.146	0.146	0.153
Participated in first and last year	0.607	0.609	0.825	0.607	0.606	0.609
Average Participation in a year	0.613	0.619	0.815	0.613	0.612	0.615
Rebalancing rate	0.599	0.600	0.716	0.599	0.599	0.601
α_1- 1 quarter	0.029	0.029	0.031	0.028	0.028	0.029
α_1- 4 quarters	0.050	0.053	0.090	0.050	0.051	0.051
α_1- 16 quarters	0.115	0.100	0.229	0.113	0.112	0.090
Equity share	0.784	0.782	0.833	0.784	0.784	0.788
Median of Fin Wealth to Wages (Stockholders)	0.886	0.889	1.234	0.886	0.887	0.887
Median of Fin Wealth to Wages (All households)	0.517	0.524	0.918	0.517	0.517	0.518
Risk Premium	0.074	0.074	0.074	0.074	0.074	0.075