Endogenous Stock Market Participation: Risk Preference and Participation Cost

Qizhou Xiong *

IWH & Otto-von-Guericke University Magdeburg

January 4, 2019

Abstract

This paper revisits the limited stock market participation puzzle and treats non-participation as an endogenous self-censoring. I employ a censored fractional response model to estimate both participation costs and risk preferences in a reduced form framework using European survey data. I show that all households face non-negligible participation costs, and the non-participants have significantly higher risk aversion than the participants. The new estimation corrects the sample selection bias in estimating risk preference with only participants and suggests an average relative risk aversion around 8.3 and an average participation cost of approximately EUR 80 for European households.

JEL code: D14 D15 G10 C24

Key words: Limited stock market participation, relative risk aversion, participation cost

---

*Halle Institute for Economic Research (IWH) and Otto-von-Guericke University Magdeburg, Germany. E-mail: qizhou.xiong@iwh-halle.de. I would like to thank Sylvain Chabe-Ferret, Stefano Colonnello, Pierre Dubois, Jean-Pierre Florens, Mattia Girotti, Michael Koetter, Pascal Lavergne, Margaret Leighton, Thierry Magnac, and all the participants of Applied Microeconometrics Workshop at TSE for insightful comments and advices. This research has received financial support from the European Research Council under the European Community's Seventh Framework Program FP7/2007-2013 grant agreement No. 295298. The usual disclaimer applies.

1 Introduction

Households rarely hold risky assets. Less than half of the population hold risky assets in countries with highly developed financial markets. In the Household Finance and Consumption Survey (HFCS), conducted by the European Central Bank, only approximately 33% of households hold risky assets. In the US, Haliassos and Bertaut (1995) show that less than 50% of households held stocks either directly or indirectly in the mid-1980s. Bricker and Windle (2014) document that only 39.1% in 2010 and 35.5% in 2013 among the surveyed households held risky assets. The participation rate is even lower in developing economies. China, for instance, only had 9.0% stock market participation according to Chinese Household Finance Survey 2011.

The non-participation in the stock market[1] is particularly puzzling within the optimal portfolio framework (Merton, 1969, 1971), which states that the share of wealth invested in risky assets is a direct reflection of the investor's relative risk aversion. If the zero holding were truly the optimal risky share for the non-participants, one needs to employ infinite relative risk aversion to reconcile such action. Therefore, empirical exercises using such a revealed preference approach often ignore the non-participants and estimate the relative risk aversion with only the households who have a positive share of risky assets (Friend and Blume, 1975; Morin and Suarez, 1983). Overlooking the non-participants in estimating relative risk aversion may lead to sample selection bias if the non-participants and participants have substantially different risk preferences.

One way to justify the non-participation is to consider that the households face a per period participation cost to stock market (Luttmer, 1999; Vissing-Jorgensen, 2002). Trading in the stock market comes with a cost due to information collection and processing, mental costs of portfolio re-balancing, and the opportunity costs from the time spent in portfolio manag-

[1] I use "stock market participation" to define positive risky assets holdings directly and indirectly. The definition of risky assets follows Brunnermeier and Nagel (2008). Financial assets such as savings bonds, deposits, pension, and insurance investment are considered to be risk-free assets; mutual funds, publicly traded stock, private business investment, managed accounts, and money owed to households are risky assets.

ing activities instead of leisure. When the optimal level of risky asset holding, determined by households' risk preference, brings the households less benefit than the participation cost, they decide not to participate at all (Attanasio et al., 2002; Vissing-Jørgensen, 2002). However, this does not necessarily mean that non-participants are always more risk averse. The non-participation can be the result of either high risk aversion, high participation cost, or both. This makes non-participation an informative action that contains a noisy message on both participation costs and the risk preference of non-participants. Building on the stock market participation cost literature, I model the non-participation as a well-informed rational choice instead of an investment mistake (Calvet et al., 2007). Therefore, I can consider the non-participation to be the result of self-censoring, which allows me to use the whole sample instead of participants only in the empirical exercise, and to answer the question of whether participants are less risk averse.

I assume that both participation costs and risk preferences are heterogeneous and correlated with household characteristics, based on substantial empirical evidence in the literature (Jianakoplos and Bernasek, 1998; Halek and Eisenhauer, 2001; Van Rooij et al., 2011). I apply a two-stage estimation to estimate the parameters of the participation cost function and relative risk aversion function using the first two waves of the HFCS. In the first stage, I apply extremal quantile regression to estimate the heterogeneous participation cost function that depends on household characteristics such as education, gender and income. In the second stage, I use the estimated participation cost function to create individual censoring thresholds and then estimate the censoring model with a maximum likelihood estimator. It is worth noting that the risky share of the financial wealth is a fractional response variable belonging in the unity interval with a large share of corner observations. I adopt the logistic transformation, as in Papke and Wooldridge (1996), and add censoring to include corner observations in the quasi-maximum likelihood estimation.

The main findings of the paper is that the participants are indeed significantly less risk averse and that we should consider higher average relative risk aversion for the representative agent. Based on standard parameter choices of the stock market, the estimated median relative risk aversion is approximately 8.35 in Europe, which is substantially higher than what was previously estimated. Moreover, the results show that the non-participants have significantly higher relative risk aversion, which confirms the downwards sample selection bias if we estimate using only the sub-sample of participants.

Comparison among the estimation methods shows the magnitude of the sample selection bias by leaving non-participants out of the estimation. For instance, the estimated relative risk aversion from applying the methods of Chiappori and Paiella (2011) and Friend and Blume (1975) to the same data are approximately 20% and 80% lower respectively. I also employ the Survey of Consumer Finance (SCF) from the U.S. to compare the estimation methods and eliminate data specificity concerns from using only European data. The SCF results show substantially higher estimated risk aversion of the censored fractional response model than the estimates of the previous revealed preference approaches, confirming the sample selection bias again from the U.S. survey data.

The estimated average participation cost in this paper is from 0.07% to 0.1% of the total financial wealth, which is consistent with the previous findings in the literature. Paiella (2007) find that the foregone gain from not investing in risky assets is from 0.7% to 3.3% of household non-durable consumption, which loosely translates to EUR 110 as a lower bound for participation cost. Attanasio and Paiella (2011) show the lower bound of participation cost to be 0.4% of nondurable consumption. The average participation cost translates into from EUR 70 to 90 in monetary term, which shows that participation cost is effective in deterring participation to stock market even at a very limited amount. Moreover, the similar level of participation costs compared with previous empirical evidence suggests that the higher estimated risk aversion is unlikely to be the mechanical consequences of an unreasonably high participation cost.

This paper makes a few substantial contributions to the literature on household finance and limited participation. First, to the best of my knowledge, this paper is the first to estimate the heterogeneous risk preferences and participation costs of non-participants simultaneously. The unique self-censoring framework allows estimation of risk aversion of the non-participants in a simple reduced form estimation without relying on a computationally costly structural model and rarely obtainable high quality panel data (Fagereng et al., 2017).

Second, this paper also provides robust empirical evidence on the representative risk preference of European households by correcting the sample selection bias and using relatively novel EU household finance survey data. The empirical exercise also provides rare evidence of the risk aversion of non-participants.

Third, the estimated functions of both participation costs and risk preference allow a more flexible heterogeneous agent setup for macroeconomic and asset pricing models. Modeling risk preference as a function of household characteristics and state variables not only offers an inherent time-varying risk aversion specification but also an additional mechanism channel to explain well-known economic phenomena such as safe asset shortage and risk premium puzzle (Gomes and Michaelides, 2008; Caballero and Farhi, 2017).

Finally, the censored fractional response model is also an improvement on the zero-inflated model in dealing with fractional variables with a large number of corner observations, and is more consistent with the underlying economic decision making process.

The rest of the paper is organized as follows. Section 2 describes the data and presents basic summary statistics. Section 3 outlines the theoretical framework. Section 4 details the empirical strategy. Section 5 shows the results and section 6 is the robustness discussion. Finally, Section 7 concludes.

2 Data

This paper uses the relatively new HFCS for the empirical exercises. HFCS is a decentralized cross sectional survey of the Eurosystem whereby participating nations conduct parallel micro-level data collection on households' finances and consumptions. The first two waves of HFCS became available in 2013 and 2016, collected from 2008–2010 and in 2014, respectively. I focus on the following 8 western European countries: Austria, Belgium, Finland, France, Germany, Italy, the Netherlands, and Spain. These countries are the major economies in the euro zone, and they have similar per capita GDP and financial development.[2] There also exists a panel part of the survey that is available in Germany, Italy, Belgium and the Netherlands. I use the panel part for the later cross-validation exercise.

The HFCS consists of three main parts: household level information, household member information and personal information. The household level part is composed of questions referring to the household as a whole. There is a main respondent for every household answering all the survey questions. The household-level questions cover households' real assets and their financing, liabilities, and credit constraints; private businesses and financial assets; inter-generational transfers and gifts and consumption/savings. The other two parts are targeted at individual household members[3]. Questions to individuals cover the following areas: employment, future pension entitlements and labor-related income (other sources of revenue are surveyed at the household level).

The demographics and personal information of household members, especially the main respondent, can also be influential in the financial decisions of a household. This paper uses the main respondent' key personal information, such as marital status, employment status and working hours. Moreover, the survey also provides personal information on other household members; e.g., age, gender, job status and education. These types of covariates may also have an impact on the household level investment decision. I focus on the household as the economic decision-making units and leave the intra household dynamics for future research.

In addition to the original data from the questionnaire, HFCS also provides critical household wealth and portfolio information, such as total net wealth, total assets, total

[2]The original survey data contain 15 countries in the first wave of HFCS. This paper does not consider the rest of the countries for the following reasons. First, the sample sizes vary throughout the different countries, for instance, only 843 observations were collected in Malta. I exclude the countries with too small a sample size. Second, the development of financial institutions differs greatly across countries. For example, the eastern European countries, Slovakia and Slovenia have a relatively low financial development index according to the financial development reports of the World Economic Forum. Third, a few countries, such as Greece and Portugal, were experiencing severe financial turmoil during the sampling period. This may have caused erratic financial behaviors by households, which prevents the recovery of the true risk attitude of households.

[3]Basic demographic information is collected for all household members, and a personal questionnaire is answered by every household member over 16 years of age.

financial wealth, and the value of a household's main residence. I take advantage of the granular details on the household finance in the HFCS and conduct an empirical exercise at the individual level at a relatively large scale. To better compare the results from the European data, this paper also employs the Survey of Consumer Finance (SCF) from the Federal Reserve Board of the U.S. to conduct the same empirical exercises. The SCF collects detailed information about family incomes, net worth, balance sheet components, credit use, and other financial outcomes, which makes it another ideal data source for studying household portfolio choices.

2.1 Summary Statistics

This section presents the summary statistics of the key variables that the estimation uses. Descriptive statistics are provided at both pooled and country levels. The baseline empirical exercise in this paper focuses on the investment decision making of the financial assets similar to the "cash on hand" in Cocco et al. (2005). Housing generally represents a large portion of households' wealth, but often lacks liquidity. Therefore, I treat the housing asset status as the background risk (Heaton and Lucas, 2000; Campbell, 2006) that can potentially influence households' risk preference.

The HFCS categorizes the financial assets into seven types: deposits, mutual funds, bonds, stocks, money owed to household, private pension and life insurance, and other assets. This paper uses the definition of risky assets from Brunnermeier and Nagel (2008) to categorize the following assets as risky: mutual funds, publicly traded stock, private business investment, managed accounts, money owed to households and others. Other financial assets such as bonds, deposits, pension and insurance investment are considered to be risk-free assets. For the baseline specification, I consider the bond as a risk-free asset as the majority of bond investment is in either state or government bonds. The sum of those types of assets divided by total financial assets gives us the risky shares – "α". Admittedly, dividing all assets into two risk classes is a relatively strong simplifying assumption. The investment decision concerning specific asset categories is also highly heterogeneous and can be explained by household characteristics. I provide extensive summary details on the household portfolio choices in both the shares and participation percentage to different asset categories in Appendix C. However, this paper does not provide further analysis to stay focused on the main decision of interest – participation in a risky asset market.

The upper panel of Table 1 illustrates the limited participation puzzle in the European context. More than 60% of the households do not hold any risky assets ($\alpha = 0$): 66% in the first wave and 63% in the second wave. The non-participation is high and consistent in Europe as it is in the U.S. Moreover, the risky share definition already includes indirect holdings of any risky asset; such a high level of non-participation signals a higher reluctance to invest in risky assets across European countries. Among participants, the risky shares are highly heterogeneous among the households, with a median level around 30%. The table also reports the stock market participation state in the eight selected countries. We can see that the stock market participation is highly heterogeneous across countries. For instance, Italy has the lowest participation in both waves, staying at approximately 15%, whereas Finland and Germany have much higher participation rates. Moreover, the participation rate remained stable between 2010 and 2014 in all the countries in the selected sample.

Table 2 reports the summary statistics of the demographic and financial variables as of 2010. Comparing the basic demographic summary statistics with the European average reported on "eurostat",[4] we can see that the survey is fairly representative. For instance, the average family size is 2.41 in the survey and the average household size in the EU-28 was 2.3 in 2010; the average total wealth in the survey is 467 thousand euros and the average wealth of the same-sized family is approximately 450 thousand euros in the western European countries. Moreover, the homeownership is 72%, which is consistent with the national average of the selected countries. To sum up, for most of the demographics, the survey results are consistent with the national account statistics in Europe, which indicates that the HFCS is indeed a representative survey. However, it is worth pointing out that the average age in the survey is significantly higher, and the proportion of retirees also much

[4] https://ec.europa.eu/eurostat/statistics-explained/index.php/Main_Page

higher, than in the national account.

3 A Simple Optimal Portfolio Choice Model

In this section, I introduce a simple portfolio choice model within the standard framework of Merton (1969), with the addition of heterogeneous participation costs to incorporate the non-participation as a rational endogenous choice. However, stock market participation or non-participation can be explained by various economic reasonings. Hong et al. (2004) shows that social interaction plays an important role in stock market participation decisions, which can be linked with the stock market awareness story proposed by Guiso and Jappelli (2005). In addition to the social network, the physical neighborhood also has impact on the stock market participation behavior, according to Brown et al. (2008). Guiso et al. (2008) points out that households may shy away from the stock market simply due to distrust of the complicated stock market trading and the negative image of financial professionals. It is nearly impossible to control for all the potential explanations. I make the admittedly strong assumption that risk preference is the primary driver of stock market participation decisions and that all the other explanations can either be incorporated in the different participation cost or considered as part of the unobservable heterogeneity.

There is plenty of empirical evidence that the ability of households to collect and process information for trading on the stock market is critical in household portfolio choice. Grinblatt et al. (2011) use Finnish data to prove that high-IQ households are more likely to participate in the stock market, and Van Rooij et al. (2011) show that financial literacy and numerical ability play a critical role in influencing stock market participation. These explanations are in line with the fixed per period participation cost as in Haliassos and Bertaut (1995), Vissing-Jørgensen and Attanasio (2003) and Paiella (2007). The less financially-literate households would have to exert more effort, thus experiencing more dis-utility, to complete the same trading task, which simply translates into a higher participation cost. I adopt the general definition of participation costs as the overall dis-utility of being involved in any stock market trading activity to explain non-participation.

3.1 Households

Consider the optimization problem of a risk-averse household, denoted by i, which lives in a two-period economy with no taxes and where all assets are liquid. Each household has an initial endowment W_{i0} at the beginning. Households have the standard borrowing constraint that they cannot invest more than their total financial wealth. At period zero, households invest their wealth and consume nothing. In period one, households gain the return of their investments and consume all wealth.

There are two types of assets in the economy – risky assets and risk-free assets. Thus households face a budget constraint as follows:

$$\mathbf{E}[W_{i1}] = W_{i0}\{1 + r_f + \alpha_i \mathbf{E}[r_m - r_f]\}, \tag{1}$$

where W_{i1} is the uncertain wealth at the end of period one; r_m and r_f are the random rate of return of the risky asset and the risk-free interest rate respectively; and α_i is the fraction of wealth invested in the risky asset, which is referred to as risky share in the rest of this paper. In a stochastic financial market identical to Merton (1969), households' expected wealth change can be expressed as follows:

$$\begin{aligned} E_0[W_{i1} - W_{i0}] &= \{r_f + \alpha_i \mathbf{E}[r_m - r_f]\}W_{i0} && (2) \\ E_o[(W_{i1} - W_{i0})^2] &= W_{i0}^2 \alpha_i^2 \sigma_m^2. && (3) \end{aligned}$$

Households derive their utility through a smooth concave utility function $U(W_i)$, and they try to maximize the total expected utility with respect to budget constraint, as described in Equations 2 and 3.

3.2 Participation Cost

Trading in the stock market and risk-free markets comes with a fixed per period cost. Households must participate in at least one type of financial market to smooth their inter-temporal consumption. There are three possible choices for each household:

1. "Complete portfolio": households participate in both risky and risk-free markets – $\alpha_i \in (0,1)$. Most of the previous literature focuses on this type of household.

2. "Risk-free only": households participate only in the risk-free market — $\alpha_i = 0$. This describes the majority of the observed households in many surveys.

3. "Risky only": households invest their entire financial wealth in the risky asset – $\alpha_i = 1$. Such behavior can be explained only by very low risk aversion and high participation costs in both markets. This type of household is not often observed in any survey data.

Normalize the participation cost to both markets to zero. Define the participation cost incurred when the household participates in the risk-free market only, rather than both markets, as $-\delta_i^s$ and the participation cost incurred when the household participates in the risky asset market only, rather than both markets, as $-\delta_i^b$. The negative participation costs in such normalization can be understood as the benefit of simplifying the investment tasks. It is also very unlikely that households would choose only between "Risk-free only" and "Risky only". This would imply that households do not have stable risk preference at all. We can consider the three possible investment choices as ordered discrete choices with respect to optimal risky shares.

3.3 Utility Optimization

Households maximize their utility with respect to the risky shares. The decision making process of each household is as follows: first, consider the optimal risk share of the household wealth under all three choices[5]; second, weigh the total expected utility of the choices; and,

[5]The optimal risky share is given automatically under the choice "risk-free only" and "risky only".

third, choose the investment choice that returns the highest total expected utility. As choices are ordered with respect to risky shares, the comparison between choice "risk-free only" and "risky only" is captured by the other two comparisons. For instance, when a household prefers "risk-free" to "complete portfolio", the choice "risky only" is dominated by the other two choices. There is no need to consider the choice switching between "risk-free only" and "risky only".

Since the choice "risky only" is seldom observed, this paper focuses on the comparison between the choices "risk-free only" and "complete portfolio". Households with "complete portfolio" maximize $\mathbf{E}[U(W_{i1})]$ with respect to α_i, and with the budget constraint of

$$\mathbf{E}[W_{i1}] = W_{i0}\{1 + r_f + \alpha_i \mathbf{E}[r_m - r_f]\}. \tag{4}$$

Taking the Taylor series expansion of $U(W_{i1})$ around W_{i0} and keeping the first two terms, the expected utility becomes:

$$\mathbf{E}[U(W_{i1})] = U(W_{i0}) + U^{'}(W_{i0})E_0[W_{i1} - W_{i0}] + \frac{1}{2}U^{''}(W_{i0})E_o[(W_{i1} - W_{i0})^2].$$

Using the results of Equations 2 and 3 to replace the expected return and volatility,

$$\mathbf{E}[U(W_{i1})] = U(W_{i0}) + U^{'}(W_{i0})W_{i0}\{r_f + \alpha_i \mathbf{E}[r_m - r_f]\} + \frac{1}{2}U^{''}(W_{i0})W_{i0}^2\alpha_i^2\sigma_m^2. \tag{5}$$

At the optimal choice of the risky share α_i^*, expected utility is maximized and the first order derivative of $\mathbf{E}(U(W_{i1}))$ is equal to zero. Taking the derivative of Equation 5 with respect to α_i gives us :

$$\alpha_i^* = -\frac{U^{'}(W_{i0})}{U^{''}(W_{i0})W_{i0}} \cdot \frac{\mathbf{E}[r_m - r_f]}{\sigma_m^2} = \frac{1}{\gamma_i} \cdot \frac{\mathbf{E}[r_m - r_f]}{\sigma_m^2} \tag{6}$$

where γ_i is the Arrow-Pratt measure of relative risk aversion, and the borrowing constraint imposes the binding constraint that $\alpha_i^* \in [0, 1]$. Equation 6 is as derived by Friend and

Blume (1975) and Chiappori and Paiella (2011).

The value of choosing "complete portfolio" is

$$\begin{aligned}\mathbf{V}_s(W_{i0}) &= U(W_{i0}) + U^{'}(W_{i0})W_{i0}\{r_f + \alpha_i^*\mathbf{E}[r_m - r_f]\} \\ &\quad + \frac{1}{2}U^{''}(W_{i0})W_{i0}^2\alpha_i^{*2}\sigma_m^2. \end{aligned} \tag{7}$$

If household chooses the portfolio "risk-free only", the total expected utility is:

$$\mathbf{V}_0(W_{i0}) = U(W_{i0}) + U^{'}(W_{i0})W_{i0}[r_f - (-\delta_i^s)] \tag{8}$$

Therefore, the utility difference between "complete portfolio" and "risk-free only" is:

$$V_s - V_0 = U^{'}(W_{i0})W_{i0}\{\alpha_i^*\mathbf{E}[r_m - r_f] - \delta_i^s + \frac{1}{2}\frac{U^{''}(W_{i0})}{U^{'}(W_{i0})}W_{i0}\alpha_i^{*2}\sigma_m^2\}$$

We can use the optimal risky share result to simplify the value difference as follows:

$$\begin{aligned}V_s - V_0 &= U^{'}(W_{i0})W_{i0}\{\alpha_i^*\mathbf{E}[r_m - r_f] - \delta_i^s - \frac{1}{2}\gamma_i\frac{1}{\gamma_i}\frac{\mathbf{E}[r_m - r_f]}{\sigma_m^2}\alpha_i^*\sigma_m^2\} \\ &= U^{'}(W_{i0})W_{i0}\{\frac{1}{2}\mathbf{E}[r_m - r_f]\alpha_i^* - \delta_i^s\} \end{aligned} \tag{9}$$

Notice that $U^{'}(W_{i0})$ and W_{i0} are all positive. Thus, households choose the "complete portfolio" when $\frac{1}{2}\mathbf{E}[r_m - r_f]\alpha_i^* - \delta_i^s \geq 0$, and choose "risk-free only" otherwise. Similarly, we can conduct the same analysis with the tradeoff between "complete portfolio" and "risky only", which would lead us to the condition that only when $\mathbf{E}[r_m - r_f](\alpha_i^*/2 + \frac{1}{2\alpha_i^*} - 1) - \delta_i^b \geq 0$, households choose the "complete portfolio" over "risky only". I provide more detailed results on this in Appendix A

3.4 Censored Fractional Response Model

After deriving the theoretical thresholds of stock market participation, we can therefore summarize households' decisions on risky share as follows:

$$\alpha_i = \begin{cases} 0 & if \quad \alpha_i^* \in (0, \underline{\alpha}_i), \\ \alpha_i^* & if \quad \alpha_i^* \in [\underline{\alpha}_i, \bar{\alpha}_i], \\ 1 & if \quad \alpha_i^* \in (\bar{\alpha}_i, 1). \end{cases} \tag{M}$$

where $\underline{\alpha}_i \equiv 2\delta_i^s/\mathbf{E}[r_m - r_f]$ and $\bar{\alpha}_i \equiv (\frac{\delta_i^b}{\mathbf{E}[r_m - r_f]} + 1) - \sqrt{(\frac{\delta_i^b}{\mathbf{E}[r_m - r_f]})^2 + \frac{2\delta_i^b}{\mathbf{E}[r_m - r_f]}}$.

Within the simple theoretical framework of this model, I consider the observed "optimal risky share" decisions as the result of individual risk preference and unobserved heterogeneity in financial decision-making. The unobserved heterogeneity comes from sources such as investment mistakes (Calvet et al., 2007), inattention to the stock market (Abel et al., 2007), and medical expenditures (Goldman and Maestas, 2013). I also assume that the stock market information is publicly available to all households. Thus the stock market return and volatility should be perceived commonly by all households. I parametrize the optimal risky share defined in Equation 6 as a linear additive function:

$$\alpha_i^* = G(X_i\theta_M + u_i) \tag{10}$$

Where X_i represents the household characteristics that correlate with the relative risk aversion and u_i is the error term as described above. Assume that u_i is normally distributed for simplicity, which will be relaxed later in the robustness discussion. For further extension of such work, we can drop the distributional assumption on the error terms and fully adopt a non-parametric approach since the extremal quantile regression do not require any functional form of unobservable heterogeneity. $G(\cdot)$ is the logistic transformation function that ensures $\alpha_i^* \in (0, 1)$. Once we estimate θ, we will be able to recover the deterministic part of risk aversion using:

$$\hat{\gamma}_i = \frac{\mathbf{E}(r_m - r_f)}{\sigma_m^2} \cdot \frac{1}{\mathbf{E}[G(X_i\hat{\theta}_M)]}. \tag{11}$$

Moreover, I also parametrize the censoring thresholds as participation functions of the household characteristics Z_i:

$$\underline{\alpha}_i = f^c(Z_i, \theta_L) \quad \& \quad \bar{\alpha}_i = f^c(Z_i, \theta_H). \tag{12}$$

I adopt the similar linear additive logistic transformation function $G(\cdot)$ for the participation cost function $f^c(\cdot)$ to ensure that the risky share thresholds are also confined in the unity interval.

This model deviates from a standard Tobit model as the dependent variable here is a fractional variable. In the fractional variables literature, there are two critical features that require special treatment: the boundedness of the variable and the observations on the boundaries. To deal with the boundedness of fractional response variables, I adopt the logistic transformation to ensure that the response variable is bounded in $(0, 1)$. Such an approach overcomes the mismatch between the support of the dependent variable and the support of the error term, which is a clear drawback in a simple OLS regression. Moreover, this also implies that the conditional expectation function is nonlinear since it maps onto a bounded interval; and its variance decreases to zero as the mean approaches either boundary. The parametric specification of α_i^* allows for the precise definition of censoring probability. Papke and Wooldridge (1996) propose a quasi parametric approach in which $\mathbf{E}[\alpha|X] = G(X\theta)$, where $G(\cdot)$ is a known function satisfying $0 < G(z) < 1$ for all $z \in \mathbf{R}$. There is no assumption on the distribution of error term. Alternatively, a fractional response model can also be estimated by beta regression models (Paolino, 2001; Ferrari and Cribari-Neto, 2004), which I test in the robustness section later. This paper uses the logistic transformation for its baseline analysis, which already offers good estimation results according to Kieschnick and McCullough (2003).

For some data, where the number of boundary observations is too sizable to be ignored, Hoff (2007), Cook, Kieschnick, and McCullough (2008) and Ospina and Ferrari (2012) introduce two-part models that treat boundary observations as discrete choices resulting from positive probability masses at boundaries. However, the decision of participating in risky asset markets depends on the benefits that one can gain from investing in risky assets, which is determined by the optimal risky share of the household. One cannot decide whether to invest in risky asset market without knowing the optimal risky share. The assumption of a two-part model reverses the decision-making process, while the censored fractional response model is consistent with the economic reasoning of the investment decision. Moreover, Fagereng, Gottlieb, and Guiso (2013) incorporate a simple Heckman selection bias correction to model the participation decision and consider participation in the stock market as a treatment effect.

The censored fractional response model is a simple and natural way to deal with both boundedness and large boundary observations that is also consistent with the underlying economic decision making. However, to achieve identification of all the parameters, I have to assume that the participation cost is deterministic of the observable household characteristics when households choose optimal risky shares. This admittedly strong assumption assigns the remaining unobservable heterogeneity from participation costs to the error terms of the risky shares. This may not be as detrimental considering that participation costs are the dis-utility of lost leisure time, which is also likely to be orthogonal to the control variables we use to estimate risk preference. The potential side effect is likely to increase the standard deviation of the error terms instead of biasing the parameter estimates. Nonetheless, I test the identification issue by cross-validation using first two waves of the panel part of the HFCS.

4 Empirical Strategy

There are two sets of parameters to identify in the model — (θ_L, θ_H) that govern the participation costs and θ_M that affects the risky shares. For notational convenience, I introduce three indicator random variables: $I_i^M = \mathbf{1}\{\alpha_i^* \in [\underline{\alpha}_i, \bar{\alpha}_i]\}$, $I_i^L = \mathbf{1}\{\alpha_i^* \in (1, \underline{\alpha}_i)\}$, and

$I_i^H = \mathbf{1}\{\alpha_i^* \in (\bar{\alpha}_i, 1)\}$. Altonji, Ichimura, and Otsu (2012) prove that θ_M and θ_L and θ_H in censoring functions are identified, when the independence assumption between covariates and errors holds and $G(\cdot)$ continuously differentiable.

In the first stage, we estimate the participation cost functions. I denote censoring probabilities $P_i^M(X_i) = \mathbf{Pr}\{I_i^M(X_i) = 1\}$, $P_i^L(X_i) = \mathbf{Pr}\{I_i^L(X_i) = 1\}$ and $P_i^H(X_i) = \mathbf{Pr}\{I_i^H(X_i) = 1\}$, which add up to one: $P_i^L(X_i) + P_i^M(X_i) + P_i^H(X_i) = 1$. The censoring probability is well defined under the baseline specification with the normal assumption on the random shocks.[6] As suggested by Altonji et al. (2012), we can first estimate θ_L and θ_H by extreme quantile regressions:

$$\hat{\theta}_L = arg \min_{\theta_L \in \Theta_L} \sum_{i=1}^{n} \rho_{\tau_n}(\alpha_i - \underline{\alpha}_i) I_i^M \quad for \quad \tau_n \to 0, \tag{13}$$

$$\hat{\theta}_H = arg \min_{\theta_H \in \Theta_H} \sum_{i=1}^{n} \rho_{\tau_n}(\alpha_i - \bar{\alpha}_i) I_i^M \quad for \quad \tau_n \to 1, \tag{14}$$

where function $\rho_\tau(v) = (\tau - \mathbf{1}(v \leq 0))v$. The properties of extreme quantile estimates are based on the asymptotic theory of extremal quantiles (Chernozhukov, 2005). I explain the intuition of using extremal quantile approximation in Appendix B.

In the second stage, we estimate θ_M by maximizing the following criterion function:

$$\ell(\theta_M; X_i, \alpha_i) = \sum_{i=1}^{N} \{I_i^L log(P_i^L) + I_i^M log(P_i^M) + I_i^H log(P_i^H)\} - \sum_{i=1}^{N} I_i^M (\alpha_i - G(X_i \theta_M))^2. \tag{15}$$

The standard asymptotic theory on extremum estimators applies here [see Newey and McFadden (1994)]. The asymptotic distribution of θ_M is equivalent to a standard extremum estimator with θ_L and θ_H known. Combining with the asymptotics of extremal quantile regression, we have asymptotic normality for all the parameters: $\hat{\theta} = (\hat{\theta}_L, \hat{\theta}_H, \hat{\theta}_M)$.

[6] However, it is also possible to use the empirical distribution of random shocks to derive the censoring probability as well. Then the estimation of the model becomes semi-parametric.

4.1 Extremal Quantile Regression

The estimation of the model starts with the extreme quantiles of the observed risky shares in the open interval $(0, 1)$. As only the observations with risky shares belonging to (0,1) enter this stage of estimation, applying the logistic transformation to the observed real risky share α_i^* simplifies the conditional extremal quantile regression to:

$$Q_{G^{-1}(\alpha_i^*)}(\tau|Z_i) = Z_i'\theta_L(\tau) \quad \& \quad Q_{G^{-1}(\alpha_i^*)}(1-\tau|Z_i) = Z_i'\theta_H(1-\tau). \tag{16}$$

We can, therefore, obtain the modified conditional extremal quantile estimators as in Equation 13 and 14. This estimator follows closely the method proposed by Chernozhukov and Fernández-Val (2011). I set the extreme quantile at $\tau = 0.05$. The rule of thumb in deciding whether to use extreme quantile regression is that $\tau \cdot N/50 \leq 20$. The sample size for the pooled extremal quantile regression is 16415, which still fits the rule of thumb with a $\tau = 0.05$. However, the rule of thumb only specifies the upper bound for using extremal quantile regression. If τ is too large, it defeats its purpose of approximating the censoring thresholds. If τ is too small, the observations below the quantile would be insufficient, which biases the estimation. Moreover, given that the average financial wealth in the data is approximately fifteen thousand euros, if one invests less than 5% of one's financial wealth in risky assets, the return will be nearly 750 euros. If such households invest in an asset with risk premium rate of return and 100 % certainty, they would earn only approximately 50 euros, which is barely enough to cover the basic transaction costs and account management fees. Applying similar logic, Chiappori and Paiella (2011) delete all the observations with less than 3.5 % of financial wealth in risky assets.

The conditional extremal quantile regression uses the sub-sampling method to obtain the unbiased asymptotic statistics. The estimation sets the sub-sample size to be 30% of the total sample for each re-sampling. The asymptotic distribution of the parameters is derived from resampling 199 times. [7]

[7] I chose this low number for the sake of reducing the computation burden. A higher number of simulation

Given the definition of participation costs in the model, the baseline specification incorporates five variables to explain the participation costs of the households: Age, gender, education, labor status, and total household income. The participation costs to financial markets come from two main channels: the information processing cost and the opportunity cost of leisure.

Age, gender, and education may be able to capture the learning cost for households. Age is an important indication of the strength of the learning ability of investors. On the one hand, the younger investors are more used to modern technology, which allows them to learn the rules and process the information more quickly; on the other hand, the older investors have more experience and connection, which potentially allows them to obtain the information and deal with the investment tasks more efficiently. Education is another obvious variable that can affect the participation cost. Education indicates people's cognitive ability and learning ability; for instance, people with higher education should have a relatively lower learning cost in understanding financial markets.

The labor status and total income of the households are indicative of the opportunity cost the households face. The labor status in this study is a dummy variable indicating whether he/she is a wage earner. The employed may be too busy to go to the trouble of trading in the financial market which leads to a higher opportunity cost for participating the risky asset markets. Meanwhile, employed people may have more exposure to financial market information that encourages them to invest. For instance, many employees hold stocks of their own companies. This implies that people obtain inside information about either their own companies or closely related companies.

4.2 Censored Fractional Response Regression

In the second stage, we use the results from extremal quantile regressions and specify the censoring probabilities to run the maximization of the criterion function in Equation 15.

will be carried out in the robustness checks.

The censoring probabilities P_i^M, P_i^L and P_i^H take the functional form as follows:

$$\begin{cases} P_i^L(X_i, Z_i) = Pr\{G(X_i\theta_M + u_i) \leq G(Z_i\hat{\theta}_L)\} = \Phi_{(0,\sigma^2)}(Z_i\hat{\theta}_L - X_i\theta_M) \\ P_i^H(X_i, Z_i) = Pr\{G(X_i\theta_M + u_i) \geq G(Z_i\hat{\theta}_H)\} = \Phi_{(0,\sigma^2)}(Z_i\hat{\theta}_H - X_i\theta_M) \\ P_i^M(X_i, Z_i) = 1 - P_i^L(X_i, Z_i) - P_i^H(X_i, Z_i) \end{cases} \tag{17}$$

where $\Phi_{(0,\sigma^2)}(\cdot)$ is the normal cumulative distribution function that centers at zero with a standard deviation of σ, which is estimated simultaneously with the rest of the parameters.

The selection of explanatory variables is based on the previous findings in the literature. The financial wealth or the total wealth of the households is always the key determinant of risk preference. The early work of Morin and Suarez (1983) includes age and wealth as the two main determinants of demand for risky assets. In a more recent work, Brunnermeier and Nagel (2008) investigate whether liquid wealth changes would induce a change in the demand for risky assets with panel data. This paper acknowledges the importance of wealth in determining the optimal risk shares. However, due to the cross-sectional data limitation, I do not interpret the correlation between relative risk aversion and wealth.

It is well-documented in the literature that demographic characteristics of households have a systematic impact on risk preference. Barsky, Juster, Kimball, and Shapiro (1997) show that age and gender partially explain the heterogeneity in risk attitude. Education is also one of the potential determinants of risk aversion. However, the previous literature does not reach a consensus about the impact of education on risk preference. Chiappori and Paiella (2011) show that education has positive impact on risk taking behavior and Christiansen, Joensen, and Rangvid (2008) determine that economists are more likely to own stocks indicating that highly educated people may have lower risk aversion. But in a study of the twins' financial behavior using the Swedish tax registry data, Calvet and Sodini (2013) conclude that the general level of education does not influence risk preferences. Moreover, Love (2010) studies the effect of marital status and children and finds that a larger family with children might take less financial risk.

Background risk, such as human capital, housing wealth, and private business, can also alter households' risk preference. For instance Cocco, Gomes, and Maenhout (2005) investigate the effect of human capital on risk aversion, and Campbell (2006) mentions that housing might be the largest background risk. Moreover, Heaton and Lucas (2000) discuss private business investment as the background risk to explain the reluctance of investing in risky financial assets. I include employment status, net housing asset and pension status to control for the background risk as well as the cross sectional survey data permits.

5 Results

This section presents the results of the empirical exercises. The baseline specification assumes that all the explanatory variables affect households the same way in all eight different countries, and the country difference is captured only by the country fixed effects. It is worth stressing that some abnormal sample observations are screened out in the data such as holding 20 euros in a mutual fund or 50 euros in a stock market. As cases of 100 % risky share are extremely rare, I will focus on the results at the lower censoring thresholds. However, I will not focus on understanding the country difference as there are often substantial institutional and cultural differences between the countries.

5.1 The Participation Costs

Extremal Quantile Regression Results of the first stage are presented in the upper panel of Table 3. I show both lower extreme quantile regression and upper extreme quantile regression results. In general, the variables explain the lower censoring thresholds much better than they do the upper censoring threshold; this is due to the limited amount of upper censored observations in our sample. Therefore the analysis of the results will focus on the lower censoring thresholds, which corresponds to the participation cost of participating in both risky and risk-free markets instead of participating in a risk-free market only.

In both waves of HFCS, most of the control variables included in the regression are significant, apart from gender. Age has a positive effect on the participation costs, which means that the older the household, the higher the participation costs. This signals that the declining learning ability due to aging is dominant in determining the participation cost. Education is a good indication of learning and information processing ability, which would translate to a negative effect on participation costs. Intuitively, people with higher education, on average, can learn the rules of risky asset markets and process the information more efficiently. However, education has surprising positive effects on participation costs. Note that the participation costs in this paper are the subjectively perceived participation costs and lost leisure. Such positive impact of education probably indicates that highly-educated people may value more the leisure time lost due to trading activity. Moreover, income is negatively correlated with participation costs indicating that higher-income households tend to have a lower participation cost proportional to their total financial wealth.

Based on the estimation results, we calculate the lower bound of risky shares, and the participation costs measured in percentage of financial wealth and monetary terms as shown in the lower panel of Table 3. From the results, we can see that the median level of lower thresholds for risky share is approximately 2% in both waves of the survey data. And the estimated thresholds do not show a high level of heterogeneity as the interquartile range of the estimated thresholds is only approximately 1%. The corresponding percentage average participation costs concerning total financial wealth are 0.07% in HFCS 2010 and 0.1% in HFCS 2014. However, most of estimations of the participation costs are expressed as a share of total consumption. Luttmer (1999) was one of the first to quantify the participation costs, and he finds the cost to be at least 3% of monthly per capita consumption. Paiella (2007) concludes that the lower bound for participation cost ranges from 0.7% to 3.3%. Attanasio and Paiella (2011) estimate the participation cost to be 0.4% of nondurable consumption. Unfortunately, the HFCS does not provide information on the total consumption, and it is hard to evaluate whether the estimates are in line with the previous findings.

To draw meaningful comparison, we turn to the monetary equivalent participation cost. I calculate the monetary cost using the estimated percentage participation cost and the financial wealth of the households. On average, households face a participation cost from

70 to 90 euros in the two waves of the survey data. These results are in line with the previous literature. For instance, Vissing-Jorgensen (2004) does a simple estimation finding that a per period stock market participation cost of approximately 55 dollars in the year 2003 prices is enough to explain the non-participation of half the nonparticipants. Mulligan and Sala-i Martin (2000) find that the participation cost is 111 dollars per year. The results of Attanasio and Paiella (2011) show that the participation cost is 72 euros per year on average. However, the median costs are much lower: the median participation cost is only 9.9 euros in HFCS 2010 and 11.7 euros in HFCS 2014. This shows that even a small amount of participation cost is effective in deterring stock market participation.

5.2 The Relative Risk Aversion

The censored fractional response model estimates the expected demand for risky shares, which leads us to the estimation of relative risk aversion using the equivalence in Equation 11. Table 4 presents the estimates of relative risk aversion in the Euro zone from the first two waves of HFCS in the lower panel of the table. We can see that the average relative risk aversion is approximately 8.8 and the median risk aversion is 8.3. The estimates are very stable between 2010 and 2014 even in the same period, the global economy was experiencing financial crisis. Compared with the previous findings, my estimates are significantly higher. For example, Chiappori and Paiella (2011) state that the average relative risk aversion for the Italian population should be 4.2 (2.5 if you do not account for the households who hold less than 6% of risky shares). Attanasio and Paiella (2011) estimate relative risk aversion using the U.S. Consumption Expenditure Survey and find that the average relative risk aversion is 1.7. Finally, Friend and Blume (1975), the first paper that estimates relative risk aversion using the revealed preference approach and risky shares, find that the relative risk aversion is approximately 2. Most importantly, many macroeconomic calibration studies consider the relative risk aversion to be between 2 and 4. Given that most of the calibration results are sensitive to risk aversion parameters, such a large dispersion between my results and the conventional level could undermine many macroeconomic and asset pricing calibration results.

Now I can also answer the question of whether non-participants are more risk averse than the participants by comparing the estimates from different results and the estimates between two groups within the data. To make sure that the difference in the estimated relative risk aversion is not just because of the different data, I apply the methods used in the previous literature to the HFCS data. Three studies are included: Friend and Blume (1975), Morin and Suarez (1983) and Chiappori and Paiella (2011). The lower part of Table 6 reports the results estimated by those methods. Figure 1 compares the four estimates in three different ways. We can see that the estimates of this paper are higher on average than are those of the other three, which implies that by ignoring the households who do not invest in risky assets, the relative risk aversion might have been under estimated. This also implies that the stock market participants have lower relative risk aversion on average. The t-test of two subgroups – the households with and without risky asset, proves that these two subgroups have different means. The inclusion of the households who do not hold risky assets also allows the estimation to capture more heterogeneity in relative risk aversion. Additionally, Table 6 also provides the estimated risk preferences across Europe. We can see that the estimates between two waves are fairly consistent, but they show substantial heterogeneity across countries as shown in Figure 2.

Moreover, I also compare the estimated risk aversion using the censored fractional response model between participants and non-participants to see whether they have different risk preferences. We can see from Table 5 that participants are indeed less risk averse than are the non-participants; however the margin is not large. The simple two-sample t-test confirms that participants have lower relative risk aversion, but it is only approximately 15% lower in both waves of HFCS. Meanwhile, we also observe that participants are significantly wealthier than are non-participants, but with a similar percentage participation costs. This implies that most of the non-participants are less wealthy households with limited financial wealth whose optimal risky shares are not sufficient to compensate for the participation costs they face.

5.3 Determinants of Relative Risk Aversion

The dependent variable in the censored fractional response model is risky share. Thus the regression is an estimation of demand for risky assets. The first part of Table 4 reports the regression results. With the assumption that asset market information is common knowledge, the heterogeneity in demand for risky assets can only be explained by heterogeneity in relative risk aversion. Therefore, with Equation 11, the impact of the explanatory variables on the relative risk aversion is opposite to that on the demand for risky assets.

First of all, the country fixed effect, albeit omitted in the table, is one of the most pronounced determinants in the regression as shown in Figure 2. This implies great cultural and institutional differences among the economically similar western European countries. The Dutch households, for instance, appear to be more risk averse than are the rest of the European countries in the estimates of both waves of HFCS. However, this may be due to the tax deductibility of mortgage payments, as discussed in the previous section. As the Dutch government partially rolled back the favorable tax scheme for mortgage, we observe a decline in risk aversion. Moreover, Spanish households seem to be the least risk averse; the consistent lower estimates in the second wave from Spanish households indicates the existence of some strong cultural factor that makes Spanish households more willing to take financial risk. The estimated risk preference is line with the time varying risk aversion because of the changing macroeconomic environment (Guiso et al., 2018). In most of the countries except for Belgium and Finland, households tend to exhibit higher level of risk aversion in the post crisis era. Especially for Belgium, we can see a drastic drop in risk aversion estimated from the observed risky shares. I have not found any legislative or regulation changes that connect to the risky asset holding incentives in Belgium.

Second, the demographics of the household do seem to drive different risk taking behavior regarding financial portfolio choices. The results suggest that age has a positive impact on the demand for risky assets, which is similar to the findings in Barsky et al. (1997). This is also consistent with the theoretical implication that people accumulate their investment in risky asset markets to earn higher permanent income in preparation for retirement. Moreover, the investment in risky assets still increases for a short period after pensionable age, as seen in the data. Marital status and family size both have significant negative impact on risky shares. This is consistent with the findings of Love (2010) that the responsibility of marriage and family members will discourage risky investment. Education shows that highly-educated people tend to have higher risky shares and thus are less risk averse, which is consistent with the findings in Christiansen et al. (2008). This is probably due either to the correlation of education and wealth or to the fact that highly educated people are better informed about the financial markets.

Third, I also test the lasting question of whether wealthier households are less risk averse. The coefficient of wealth suggests a positive answer to the question. From the results in Table 4, wealthier households tend to hold higher risk shares and be less risk averse. However, cross sectional data does not have the capacity of answering why they are less risk averse. Using panel data, Brunnermeier and Nagel (2008) find that wealth change does not change risk preference, which, in turn, signals that either wealthier households are born wealthy and are less risk averse, or less risk averse people are more likely to be wealthy. Labor income risk is one of the important background risks for the households. When people retire they no longer have labor income; thus, they no longer have any such background risk. However, the results show that such status switch does not change households' risk attitudes.

Finally, housing is usually a large part of household wealth, which imposes background risk for the households. Studies like Chetty et al. (2017) and Cocco (2004) emphasize the importance of housing in affecting portfolio choice. The results imply that net housing value has a positive impact on risky shares, which means that, the greater the housing asset of the household, the less risk averse it is. This corresponds to the similar findings in Chetty et al. (2017) that higher home equity is associated with higher stock shares in financial assets. However, the tenure status of the main residence, has a negative impact on risky shares. One possible explanation is that households view the housing investment as risky, so that they do not take on more risks in the financial market. The variable other property indicates whether the household owns property apart from the main residence. This is an indication of whether the household is using the real estate market as an investment vehicle for its portfolio. If the coefficient is positive, it means that households are more likely to consider the real estate as risk-free assets; if the coefficient is negative, households intend to treat the real estate as risky assets. The positive sign of the coefficient shows that households are less risk averse when they hold other properties and are more likely to view real estate as being less risky than are the risky financial assets.

5.4 Estimation with the Panel Part

One of the main weakness of the two stage estimation applied in the previous section is that the deterministic participation cost function is estimated from the same data sample. Moreover, the estimation of participation cost relies essentially on the extreme value distribution of the risky shares of the households who participate in the stock market. To deal with the potential identification issue, we exploit the panel part of HFCS in a few selected countries and use the first wave to estimate the participation cost and the second wave for the censored fractional response model to avoid the identification issue arising from using only one data sample. This is, in a way, similar to using the lagged variable as the instrument variable. The only assumption we need to make on the participation cost is that the parameters of the linear function remain constant, which is an improvement over the baseline identification.

I report the estimation results from such a cross validation approach using the panel part of the HFCS in Table 7. Among the eight countries in the baseline estimation, only Belgium, Germany, Italy, and the Netherlands conducted the panel part of the survey. As can be seen, using either lagged or forward estimated participation cost function does not change significantly the results of the relative risk aversion estimation both at the pooled level and at the country level. The two stage estimation process in the baseline model seems to create enough separation between participants and non-participants, that even using the same cross-section does not create an identification issue.

5.5 Evidence from SCF

For more efficient comparison with the previous results, I also apply the censored fractional response model estimation to the Survey of Consumer Finance from 1995 to 2016[8]. I adopt the risky asset categorization as in Bertaut (1998) to calculate the risky share and define the participation status. In Table 8, I present the estimated relative risk aversion and participation costs with the comparison between the participating groups and among methods in the latest SCF survey. As can be seen, the Americans are less risk averse and also face less participation costs on average. The self-censoring approach seems to correct substantially the sample selection bias of leaving non-participants out, as the estimated average risk aversion is more than 80% higher than are the previously estimated results from only participants. We can also see a much more significant difference in estimated risk aversion between the participants and the non-participants. However, if we split the groups, the estimated risk aversion of the participants from the new method are very close to the previous results, which confirms that the higher average estimated risk aversion can be attributed to the inclusion of the more risk averse non-participants. The results in Table 8 also confirm the clear financial wealth and total wealth gap between the participants and non-participants in the U.S. as in Europe. However, it is worth noticing that the participation cost in monetary term is much smaller in the U.S. compared with the previous estimates using U.S. data and the estimates from Europe. This could be an indication of higher financial market and service development in the U.S. Although the results from SCF 2016 are not perfectly in line with what we have found in the European results, the clear main message is that the non-participants are indeed more risk averse than are the participants; and leaving them out can create substantial sample selection bias.

Additionally, I apply the estimation methods to all the seven waves of SCF to see the consistency of the estimated risk aversion and potentially capture the time varying feature of the risk preference. In Figure 3, we can see the evolution of estimated relative risk aversion

[8]SCF collects survey data every three years. The two decades of data I use covers 7 waves in total.

in SCF using the repeated cross section surveys. First, consistent disparity exists between the participants and the non-participants, confirming our previous results in European data and SCF 2016. Second, the average level of relative risk aversion varies over time, which provides additional evidence for the time-varying risk aversion as in Guiso et al. (2018). Third, there is no clear pattern in average relative risk aversion except for much higher estimates in SCF 1995. However, it is evident that the average risk aversion across all groups is significantly higher after the Great Recession (in 2010), and return to previous levels as the macroeconomic environment stabilizes and recovers. This indicates that households' risk preference is time-varying and affected by recent experience in the stock market and the macroeconomic environment. Interestingly, the participants seem to become significantly less risk averse after the recovery from the Great Recession.

6 Robustness of the Results

There are a few assumptions, and empirical setting in the baseline specification can be relaxed to test the robustness of the results found in the previous section. This section discusses a few factors that may affect the estimates of relative risk aversion to see whether the estimated results of the censored fractional response model is sensitive to changes in those factors. I focus the robustness checks on the HFCS 2010 survey data, as Table 9 and Table 10 show the estimation results of different specifications.

6.1 Definition of Risky Assets

The definition of risky assets plays an obvious role in the estimation results of demand for risky assets. In the baseline specification, the bond is categorized as a safe asset, since most of the bonds are state or national government bonds. In a survey data with more details on bonds, the government bonds are considered normally as risk-free assets and corporate bonds are considered as risky assets. Moreover, some studies adopt a narrower definition of risky assets that includes only stocks and mutual funds.

This paper considers two alternative definitions of risky assets to study whether the high risk aversion is affected by the definition. The first alternative definition treats the bonds as risky assets instead of a risk-free asset, which systematically increases the risky shares for the households who hold bonds in their portfolio. The second alternative definition considers only the stocks and mutual funds as risky assets, which lowers systematically the risky shares for the households who hold other types of risky assets in the baseline specification, such as money owed to the household and private business investment.

In the second and third columns of Table 9, a different definition of risky assets does not have much effect on the coefficients of extremal quantile regression and censored fractional response regression. Compared with the baseline results, all the coefficients stay consistent in two alternative definitions. The second and third column of Table 10 shows the summary statistics of the censoring thresholds, participation costs, and estimated relative risk aversion. In these tables, only the lower censoring thresholds and participation costs are reported since the upper censoring thresholds estimation is insignificant due to the lack of observations. The censoring thresholds and participation costs remain consistent with the baseline results despite the different definition of risky assets. In particular, the results in the second column are very close to the baseline results. The estimated relative risk aversion shows the expected deviation from the baseline results. Adding bonds to the risky assets category has made the estimated relative risk aversion lower, but the difference is not very drastic given that, on average, only 5.3 % of the households hold bonds and only 6.6 % of the financial wealth is invested in bonds. The narrow definition of risky asset leads to higher estimated risk aversion, as predicted by the theoretical link between risk attitude and risk shares.

In summary, the definition of risky assets does not affect the coefficients of the estimation, but it has a systematic impact on the estimate relative risk aversion. Treating the bonds as a risk-free asset in the baseline specification does not matter as much as does considering only stocks and mutual funds as risky assets.

6.2 Extremal Quantiles

A high participation cost can increase the probability of being censored and leads to the conclusion of high risk aversion. In the baseline specification, the extremal quantile that approximate the censoring thresholds for each household are 5%, which corresponds to the minimal risky share of the average financial wealth that covers the transaction cost once a year. The fourth and fifth columns of Table 9 and Table 10 show the results of estimation when the extremal quantiles are set to be 2% and 10%. The signs of the coefficients in extremal quantile regressions remain consistent with the baseline results, but the sign changes especially in the case of extremal quantile being 2%. It is expected that the level of extremal quantiles changes the censoring thresholds and participation costs. However, the censored fractional response model estimation results are not influenced substantially by the extremal quantile setting.

The estimated risk aversion under the extremal quantile of 2% and 10% differ mildly from the baseline results. In the estimation, using a higher extremal quantile to approximate the censoring threshold will force the distribution of risky share lean towards the censoring threshold, and higher relative risk aversion is estimated. In general, the level of extremal quantiles is not the dominant factor that determines the high relative risk aversion found in the baseline results.

6.3 Estimation with Beta Distribution

Recall that in the baseline specification, the random shock to relative risk aversion is assumed to be additive to the linear index within the logistic transformation function, and to be normally distributed (Equation 17). Another common choice of distribution in the context of fractional variables is beta distribution. The last column of Table 9 and Table 10 shows the result of using beta distribution, instead of normal distribution, to model the censoring probabilities.[9] Using beta distribution gives us a significantly higher estimate

[9]The detailed specification can be found in Appendix D

of risk aversion. This is probably due to the asymmetric and flexible structure of beta distribution, which allows the beta distribution to better describe the observed empirical distribution. With a large amount of zero observations in the original data, the flexibility of beta distribution makes the distribution tilt more towards zero compare with the normal distribution.

Another reason that one may want to use Beta distribution is that it performs significantly better in predicting the censoring in simulated results. In the baseline specification, the symmetry of normal distribution cannot capture the skewness towards zero of the risky shares. This leads to the unsatisfactory prediction of 33 % of the households being censored as oppose to 66% in the data. However, when we use Beta distribution, there will be 55% of the households being censored in the simulation. I still keep the normality assumption in the baseline for simplicity, and the fact that asymmetric error term distribution is rarely used in similar context.

6.4 Perception of Market Return and Volatility

The estimated relative risk aversion has a linear relationship with the market price of risk in risky asset markets as in Equation 11. Households' expectation of risky asset return and volatility affect the estimated relative risk aversion both directly and linearly. Since the definition of a risky asset is not limited in stock markets only, the baseline specification of market return and volatility is not the observed long term stock market return and volatility. The market expectation with 8% risk premium and 20% market volatility is an optimistic opinion of the risky asset market. Chiappori and Paiella (2011) use a more pessimistic calibration of the risky asset market, with risk premium being 4% and market volatility being 20%. If we consider that other types of risky assets have market return and volatility that are similar to those of the stock market, the market price of risk has risk premium as 7% and market volatility as 23.4%, which is the post-second world war long term average stock market performance in the developed countries. Fernandez et al. (2013) provide an evaluation of the subjective expectation of market risk premium in many countries around the world. The European countries on average have market risk premium of approximately 6%. Meanwhile, the European region is known for relatively low market volatility, which means that a market volatility of 20% is likely to be a reasonable expectation.

Table 11 shows the estimated relative risk aversion with a different expectation of risky asset markets. It is evident that the perception of market return and volatility has a big impact on the results. With the pessimistic expectation of the risky asset markets, the mean risk aversion becomes close to the results of Chiappori and Paiella (2011), but the median is still much higher. Moreover, the substantial difference between participants and non-participants persists. If one tends to believe that households have perfect information about the risky assets market, the results with the post second world war developed country average are more convincing. In all the cases, the estimated relative risk aversion in this paper is still significantly higher than what was estimated in the previous literature.

7 Concluding Remarks

This paper estimates the heterogeneous relative risk aversion and participation costs under the censored fractional response model framework using data about the observed household portfolios and characteristics. The estimated results show that the risk aversion is substantially higher for non-participants and suggests a higher average risk aversion for the representative agent. Such a result is unlikely to be due to an unreasonably high participation cost, as the average participation cost in this paper falls below the range of the previous findings in monetary terms. The comparison among the methods in the survey data from both the EU and U.S. shows the magnitude of leaving non-participants out in estimating risk preference within a revealed preference framework. The robustness checks confirm that the results are not sensitive to specific empirical specifications.

To illustrate how a higher relative risk aversion parameter configuration can lead to very different implications of some well-known macroeconomic calibration studies, I re-calibrate the model of Barro (2006) with the results of this paper. I take the estimated relative risk aversion using the post-second world war developed country average performance for risky asset markets. When the relative risk aversion is 5.6, and rare disaster probability is 0.82%, Barro's model predicts the risk premium to be 6.5% and risk-free rate to be 1.6%. This is very close to the observed results from the data and has a more reasonable probability of rare disaster compared with 1.7% from the original calibration.

Moreover, the estimated functions of both participation costs and risk preference allow a more flexible heterogeneous agent setup for macroeconomic and asset pricing models. Modeling risk preference as a function of household characteristics and state variables offers not only an inherent time-varying risk aversion specification but also an additional mechanism channel to explain well-known economic phenomena such as safe asset shortage and risk premium puzzle.

However, this paper only uses cross-sectional data to study the risk preference and participation costs. It would be very interesting to apply this model in a dynamic setting with a panel dataset. The estimation results show strong country specific heterogeneity in investment behavior and risk attitude. With more information on the country-level characteristics of the economic and legislative institution, we would be able to understand such heterogeneity better.

References

Abel, A. B., J. C. Eberly, and S. Panageas (2007). Optimal inattention to the stock market. *American Economic Review 97*(2), 244–249.

Altonji, J. G., H. Ichimura, and T. Otsu (2012). Estimating derivatives in nonseparable models with limited dependent variables. *Econometrica 80*(4), 1701–1719.

Attanasio, O., J. Banks, and S. Tanner (2002). Assets holding and consumption volatility. *Journal of Political Economy 4*(110), 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*(2), 322–343.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics 121*(3), 823–866.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 537–579.

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.

Bricker, Dettling, H. H. M. S. T. and Windle (2014). Changes in u.s. family finances from 2010 to 2013: Evidence from the survey of consumer finances. Technical report, Board of Governors of the Federal Reserve System.

Brown, J. R., Z. Ivković, P. A. Smith, and S. Weisbenner (2008). Neighbors matter: Causal community effects and stock market participation. *The Journal of Finance 63*(3), 1509–1531.

Brunnermeier, M. K. and S. Nagel (2008). Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals' asset allocation. *The American Economic Review 98*(3), 713–736.

Caballero, R. and E. Farhi (2017). The safety trap. *The Review of Economic Studies 85*(1), 223–274.

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*(5), 707–747.

Calvet, L. E. and P. Sodini (2013). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *The Journal of Finance*.

Campbell, J. Y. (2006). Household finance. *The Journal of Finance 61*(4), 1553–1604.

Chernozhukov, V. (2005). Extremal quantile regression. *Annals of Statistics 33*(2), 806–839.

Chernozhukov, V. and I. Fernández-Val (2011). Inference for extremal conditional quantile models, with an application to market and birthweight risks. *The Review of Economic Studies 78*(2), 559–589.

Chetty, R., L. Sándor, and A. Szeidl (2017). The effect of housing on portfolio choice. *The Journal of Finance 72*(3), 1171–1212.

Chiappori, P.-A. and M. Paiella (2011). Relative risk aversion is constant: Evidence from panel data. *Journal of the European Economic Association 9*(6), 1021–1052.

Christiansen, C., J. S. Joensen, and J. Rangvid (2008). Are economists more likely to hold stocks? *Review of Finance 12*(3), 465–496.

Cocco, J. F. (2004). Portfolio choice in the presence of housing. *The Review of Financial Studies 18*(2), 535–567.

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.

Cohn, R. A., W. G. Lewellen, R. C. Lease, and G. G. Schlarbaum (1975). Individual investor risk aversion and investment portfolio composition. *The Journal of Finance 30*(2), 605–620.

Cook, D. O., R. Kieschnick, and B. McCullough (2008). Regression analysis of proportions in finance with self selection. *Journal of Empirical Finance 15*, 860–867.

Fagereng, A., C. Gottlieb, and L. Guiso (2013). Asset market participation and portfolio choice over the life-cycle. Technical report.

Fagereng, A., C. Gottlieb, and L. Guiso (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance 72*(2), 705–750.

Fernandez, P., J. Aguirreamalloa, and L. C. Avendaño (2013). Market risk premium used in 82 countries in 2012: a survey with 7,192 answers. *Available at SSRN 2084213*.

Ferrari, S. and F. Cribari-Neto (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics 31*(7), 799–815.

Friend, I. and M. E. Blume (1975). The demand for risky assets. *The American Economic Review 65*(5), 900–922.

Goldman, D. and N. Maestas (2013). Medical expenditure risk and household portfolio choice. *Journal of Applied Econometrics 28*(4), 527–550.

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.

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

Guiso, L. and T. Jappelli (2005). Awareness and stock market participation. *Review of Finance 9*(4), 537–567.

Guiso, L., P. Sapienza, and L. Zingales (2008). Trusting the stock market. *the Journal of Finance 63*(6), 2557–2600.

Guiso, L., P. Sapienza, and L. Zingales (2018). Time varying risk aversion. *Journal of Financial Economics 128*(3), 403–421.

Halek, M. and J. G. Eisenhauer (2001). Demography of risk aversion. *Journal of Risk and Insurance*, 1–24.

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

Heaton, J. and D. Lucas (2000). Portfolio choice in the presence of background risk. *The Economic Journal 110*(460), 1–26.

Hoff, A. (2007). Second stage dea: Comparison of approaches for modelling the dea score. *European Journal of Operational Research 181*(1), 425–435.

Hong, H., J. D. Kubik, and J. C. Stein (2004). Social interaction and stock-market participation. *The journal of finance 59*(1), 137–163.

Jianakoplos, N. A. and A. Bernasek (1998). Are women more risk averse? *Economic inquiry 36*(4), 620–630.

Kieschnick, R. and B. D. McCullough (2003). Regression analysis of variates observed on (0, 1): percentages, proportions and fractions. *Statistical Modelling 3*(3), 193–213.

Love, D. A. (2010). The effects of marital status and children on savings and portfolio choice. *Review of Financial Studies 23*(1), 385–432.

Luttmer, E. G. (1999). What level of fixed costs can reconcile consumption and stock returns? *Journal of Political Economy 107*(5), 969–997.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and statistics 51*(3), 247–257.

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of economic theory 3*(4), 373–413.

Morin, Roger, A. and A. F. Suarez (1983). Risk aversion revisited. *The Journal of Finance 38*(4), 1201–1216.

Mulligan, C. B. and X. Sala-i Martin (2000). Extensive margins and the demand for money at low interest rates. *Journal of Political Economy 108*(5), 961–991.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics 4*, 2111–2245.

Ospina, R. and S. L. Ferrari (2012). A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis 56*(6), 1609–1623.

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

Paolino, P. (2001). Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Analysis 9*(4), 325–346.

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics 11*(6), 619–632.

Van Rooij, M., A. Lusardi, and R. Alessie (2011). Financial literacy and stock market participation. *Journal of Financial Economics 101*(2), 449–472.

Vissing-Jørgensen, A. (2002). Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy 110*(4), 825–853.

Vissing-Jorgensen, A. (2002). Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures. Technical report, National Bureau of Economic Research.

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. P. Attanasio (2003). Stock-market participation, intertemporal substitution, and risk-aversion. *The American Economic Review 93*(2), 383–391.

Table 1: Stock Market Participation In Europe

Risky shares α:

		$\alpha = 0$		$\alpha = 1$		$\alpha \in (0,1)$
Obs.(2010)		32174		86		16415
Percentage		66.1%		0.2%		33.7%
	Mean	Min	25%	Median	75%	Max
	0.364	0.000	0.104	0.286	0.588	1.000
		$\alpha = 0$		$\alpha = 1$		$\alpha \in (0,1)$
Obs.(2014)		30184		110		17601
Percentage		63.1%		0.2%		36.7%
	Mean	Min	25%	Median	75%	Max
	0.378	0.000	0.107	0.304	0.614	1.000

Stock Market Participation by Country:

Country	Obs.	$\alpha \in (0,1)$	$\alpha = 0$	$\alpha = 1$	$\bar{\alpha}$	$sd(\alpha)$
HFCS 2010:						
AT	2327	0.208	0.787	0.004	0.357	0.300
BE	2316	0.371	0.627	0.002	0.391	0.309
DE	3490	0.441	0.554	0.005	0.346	0.290
ES	6023	0.337	0.659	0.004	0.519	0.320
FI	10989	0.433	0.567	0.000	0.358	0.303
FR	14916	0.346	0.653	0.001	0.319	0.275
IT	7326	0.156	0.842	0.002	0.413	0.278
NL	1288	0.337	0.659	0.004	0.264	0.253
HFCS 2014:						
AT	2997	0.194	0.803	0.004	0.365	0.295
BE	2234	0.364	0.635	0.001	0.465	0.311
DE	4433	0.456	0.538	0.006	0.336	0.296
ES	6084	0.401	0.596	0.004	0.528	0.326
FI	11008	0.538	0.460	0.002	0.363	0.293
FR	11739	0.383	0.616	0.000	0.310	0.276
IT	8156	0.117	0.881	0.002	0.457	0.297
NL	1244	0.303	0.695	0.002	0.305	0.286

This table reports the stock market participation in Europe: aggregate and country level. The upper panel shows the participation percentage and the summary statistics conditional on participation. The Lower panel shows the country level participation shares and the mean and standard deviation of risky shares conditional on participation (in column 5 and 6)

Table 2: Summary Statistics of HFCS

Summary Statistics of HFCS 2010

Variables	Mean	Min	p(25)	Median	P(75)	Max
age	54.63	16	42	55	67	85
gender	1.44	1	1	1	2	2
education	2.75	1	2	3	4	4
marital status	0.58	0	0	1	1	1
family size	2.41	1	1	2	3	15
retiree	0.34	0	0	0	1	1
employment	0.52	0	0	1	1	1
wealth(1000 €)	467.3	−1,143.2	41.3	192.6	411.0	401,119.6
financial wealth(1000 €)	103.2	0	3.4	15.3	56.1	54,444.6
total income(1000 €)	51.2.	−449.2	20.6	35.8	60.1	8,760.3
labor income(1000 €)	30.4	−138.5	0.00	18.5	44.0	2,870.0
housing tenure	0.72	0	0	1	1	1
net housing value(1000 €)	162.8	−611.5	0	120.0	230.0	8,000.0
own other property	0.36	0	0	0	1	1
public pension	0.83	0.00	1.00	1.00	1.00	1.00

Summary Statistics of HFCS 2014

Variables	Mean	Min	p(25)	Median	P(75)	Max
age	55.18	16	43	56	68	85
gender	1.46	1	1	1	2	2
education	2.92	1	2	3	4	4
marital status	0.57	0.00	0.00	1.00	1.00	1.00
family size	2.40	1	1	2	3	16
retiree	0.33	0	0	0	1	1
employment	0.52	0	0	1	1	1
wealth(1000 €)	527.4	−57,203.1	45.0	192.1	417.2	215,087.6
financial wealth(1000 €)	147.2	0.0	3.9	18.3	66.1	101,139.6
total income(1000 €)	57.8	−9,984.9	22.3	40.0	68.9	12,722.1
labor income(1000 €)	34.8	−87.8	0.0	20.4	50.7	4,330.0
housing tenure	0.72	0	0	1	1	1
net housing value(1000 €)	168.7	−58,497.5	0	120.0	232.0	10,485.6
own other property	0.35	0	0	0	1	1
public pension	0.87	0.00	1.00	1.00	1.00	1.00

I report the summary statistics of the two waves of HFCS in this table. All nominal values are reported in 1000 € terms. The gender dummy equals 2 for female head of households and 1 for male head of households. Education follows the standard internation defintion of educational attainment scale from primary to tetiery education.

Table 3: Pooled Euro Zone Participation Cost Estimation Results

Extremal Quantile Regression:

	(Intercept)	Age	Gender	Education	Employ	Income
HFCS 2010:						
$\hat{\theta}_L$	-2.922	0.004	-0.112	0.104	-0.167	-0.138
s.e.	(0.430)***	(0.002)***	(0.060)*	(0.029)***	(0.084)**	(0.035)***
$\hat{\theta}_H$	3.872	0.002	0.015	0.016	-0.496	-0.106
s.e.	(1.026)***	(0.005)	(0.112)	(0.063)	(0.145)***	(0.094)
HFCS 2014:						
$\hat{\theta}_L$	-3.06	0.01	-0.00	0.11	-0.05	-0.17
s.e.	(0.63)***	(0.00)***	(0.07)	(0.05)***	(0.10)	(0.06)***
$\hat{\theta}_H$	4.24	0.02	-0.18	-0.09	-0.41	-0.17
s.e.	(0.38)***	(0.00)***	(0.06)***	(0.04)**	(0.09)***	(0.04)***

Participation Costs Predictions:

	Mean	Min	25%	Median	75%	Max
HFCS 2010:						
$L(\hat{X})$	0.01674	0.00684	0.01442	0.01616	0.01860	0.09165
$\hat{\delta^s}$	0.00067	0.00027	0.00058	0.00065	0.00074	0.00367
Cost in €	69.7	0.0	2.3	9.9	36.7	27003.6
$H(\hat{X})$	0.9258	0.8522	0.9065	0.9184	0.9465	0.9842
$\hat{\delta^b}$	0.00013	0.00001	0.00006	0.00014	0.00019	0.00051
Cost in €	15.9	0.0	0.34	1.7	7.1	11818.1
HFCS 2014:						
$L(\hat{X})$	0.02	0.01	0.01	0.02	0.02	0.12
$\hat{\delta^s}$	0.001	0.0003	0.001	0.001	0.001	0.005
Cost in €	88.27	0.00	2.48	11.69	42.45	43,079.42
$H(\hat{X})$	0.93	0.81	0.91	0.93	0.96	1.00
$\hat{\delta^b}$	0.0001	0.0000	0.0000	0.0001	0.0002	0.001
Cost in €	21.96	0.00	0.23	1.47	7.08	88,133.67

The extreme quantile is set to be $\tau = 0.05$. s.e. is the simulated pseudo standard errors by the extremal quantile estimation. In the second panel, $L(\hat{X})$ and $H(\hat{X})$ are the censoring thresholds near zero and near one; $\hat{\delta^s}$ and $\hat{\delta^b}$ are the participation cost in percentage of total financial wealth. Cost in € are the participation costs in euro computed using $\hat{\delta^s}$ and $\hat{\delta^b}$ and household total financial wealth.

Table 4: Pooled Euro Zone RRA Estimation Results

Censored Fractional Response Model (2010):

Coefficients:	Age	Female	Educ	Marital	Family	Retiree
	0.002	-0.172	0.197	-0.217	-0.085	0.011
	(0.001)**	(0.012)***	(0.006)***	(0.014)***	(0.006)***	(0.024)
Coefficients:	Employ	Wealth	Income	Residence	Housing	Othprop
	-0.260	0.016	0.073	-0.383	0.015	0.136
	(0.021)***	(0.002)***	(0.008)***	(0.028)***	(0.002)***	(0.012)***

Country fixed effect: Yes   Pseudo R^2: 0.44

Censored Fractional Response Model (2014):

Coefficients:	Age	Female	Educ	Marital	Family	Retiree
	0.009	-0.102	0.180	-0.066	-0.031	-0.001
	(0.001)***	(0.011)***	(0.007)***	(0.012)***	(0.005)***	(0.007)
Coefficients:	Employ	Wealth	Income	Residence	Housing	Othprop
	-0.217	0.019	0.041	-0.215	0.004	0.142
	(0.014)***	(0.002)***	(0.007)***	(0.030)***	(0.002)**	(0.012)***

Country fixed effect: Yes   Pseudo R^2: 0.41

Relative Risk Aversion Estimates:

	Mean	Min	25%	Median	75%	Max
HFCS 2010	8.82	2.86	6.72	8.33	10.45	55.23
HFCS 2014	8.71	2.78	6.47	8.37	10.65	33.15

This table reports the results of the censored fractional response model with the participation cost function regarded as exogenously given. The first two panels show the correlations between risky shares and the observable household characteristics with country fixed effect. In the bottom part, we briefly report the summary statistics of the relative risk aversion imputed by the fitted value of the optimal risky share, given that the common market perception is 0.08 risk premium and 20% market volatility.

Table 5: Are Non-participants More Risk Averse?

HFCS 2010 Results:

	Min	25%	Median	Mean	75%	Max
Estimate RRA						
Participants	2.936	6.467	7.839	8.103	9.477	38.742
Non-participants	3.074	7.022	8.599	9.105	10.800	43.174
Financial Wealth 1000 €						
Participants	0	21	60	247	166	54444
Non-participants	0	2	7	29	24	24907
Total Wealth 1000 €						
Participants	-1143	166	370	902	768	401119
Non-participants	-598	16	134	244	281	80577
Participation Costs %						
Participants	0.0286	0.0576	0.0636	0.0661	0.0736	0.3081
Non-participants	0.0326	0.0590	0.0661	0.0681	0.0749	0.3349
Participation Costs €						
Participants	0.1	13.7	38.9	158.9	109.4	28150.4
Non-participants	0.0	1.2	4.7	19.7	16.1	18680.9

HFCS 2014 Results:

	Min	25%	Median	Mean	75%	Max
Estimate RRA						
Participants	2.760	5.985	7.574	7.908	9.677	25.339
Non-participants	3.034	6.851	8.785	9.143	11.169	32.515
Financial Wealth 1000 €						
Participants	0	23	65	334	180	101139
Non-participants	0	2	8	34	26	65315
Total Wealth 1000 €						
Participants	-57203	170	372	993	771	114829
Non-participants	-1509	15	124	254	270	215087
Participation Costs %						
Participants	0.0287	0.0584	0.0658	0.0680	0.0758	0.4620
Non-participants	0.0275	0.0613	0.0695	0.0733	0.0804	0.4792
Participation Costs €						
Participants	0.2	15.2	43.9	212.0	122.2	45625.3
Non-participants	0.0	1.1	5.3	23.3	18.6	27718.5

This table reports the comparison between the participants in non-participants regarding their estimated risk preferece, financial wealth, total wealth, and estimated participation costs. The relative risk aversion is computed from equation $r_h = \frac{\mathbf{E}(r_m - r_f)}{\sigma_m^2} \frac{1}{G(X\hat{\theta})}$. The market risk premium is set to be 8%, and market volatility is set to be $(0.20)^2$.

Table 6: Relative Risk Aversion Imputation

Country Level Results:

	Mean	Min	25%	Median	75%	Max
HFCS 2010:						
Austria	9.32	4.75	7.66	8.80	9.32	37.63
Belgium	7.75	3.97	6.34	7.45	8.79	20.15
Germany	8.70	4.29	7.10	8.30	9.81	27.22
Spain	4.73	2.85	3.97	4.60	5.27	15.78
Finland	8.15	3.97	6.78	7.77	9.05	28.86
France	10.84	4.69	8.74	10.53	12.39	30.21
Italy	8.16	3.74	6.93	8.01	9.09	24.76
Netherlands	14.7	7.09	11.01	13.57	17.06	54.84
HFCS 2014:						
Austria	9.42	4.98	7.64	9.15	10.90	22.80
Belgium	5.19	3.22	4.45	5.08	5.77	10.30
Germany	9.95	4.89	7.96	9.47	11.50	28.60
Spain	4.40	2.76	3.75	4.31	4.83	11.00
Finland	7.82	4.09	6.48	7.43	8.75	20.10
France	11.30	4.81	9.39	11.20	13.00	28.20
Italy	8.49	4.50	7.33	8.33	9.40	20.20
Netherlands	12.30	6.22	9.77	11.80	14.10	31.90

Friend and Blume 1975 method:

	Mean	Min	25%	Median	75%	Max
HFCS 2010	5.53	4.64	5.06	5.61	5.88	7.35
HFCS 2014	4.86	4.03	4.37	4.94	5.21	6.80

Morin and Suarez 1983 method:

	Mean	Min	25%	Median	75%	Max
HFCS 2010	4.98	3.40	4.73	4.96	5.20	83.48
HFCS 2014	4.86	3.56	4.48	4.81	5.19	17.60

Chiappori and Paiella 2011 method:

	Mean	Min	25%	Median	75%	Max
HFCS 2010	6.90	5.11	6.47	6.89	7.33	10.04
HFCS 2014	6.73	5.20	5.98	6.82	7.34	9.56

This table reports the country level estiamte risk aversion and compares etimated risk preference across different methods. The relative risk aversion is computed from equation $r_h = \frac{\mathbf{E}(r_m - r_f)}{\sigma_m^2} \frac{1}{G(X\hat{\theta})}$. The market risk premium is set to be 8%, and market volatility is set to be $(0.20)^2$.

Table 7: Relative Risk Aversion Cross Validation

RRA (2014) with Participation Costs (2010)

	Mean	Min	25%	Median	75%	Max
Baseline	8.76	3.23	6.96	8.44	10.10	32.00
Cross Validation	8.71	3.23	6.94	8.39	10.00	32.10

RRA (2010) with Participation Costs (2014)

	Mean	Min	25%	Median	75%	Max
Baseline	8.82	2.86	6.72	8.33	10.45	55.23
Cross Validation	8.77	2.94	6.81	8.29	10.39	43.28

Country Level Results 2014:

	Mean	Min	25%	Median	75%	Max
HFCS 2014:						
Belgium (B)	5.20	3.23	4.47	5.09	5.78	10.20
Belgium (C)	5.18	3.23	4.46	5.08	5.76	10.20
Germany (B)	9.95	4.90	7.97	9.49	11.50	28.70
Germany (C)	9.92	4.89	7.96	9.46	11.40	28.10
Italy (B)	8.53	4.51	7.37	8.38	9.45	20.20
Italy (C)	8.46	4.49	7.32	8.31	9.36	20.10
Netherlands (B)	12.40	6.24	9.85	11.90	14.20	32.00
Netherlands (C)	12.40	6.25	9.82	11.90	14.20	32.10
HFCS 2010:						
Belgium (B)	7.65	4.15	6.43	7.39	8.63	17.50
Belgium (C)	7.68	4.13	6.44	7.41	8.68	17.60
Germany (B)	8.60	4.61	7.12	8.19	9.62	21.10
Germany (C)	8.64	4.58	7.13	8.23	9.68	21.80
Italy (B)	8.07	4.01	6.92	7.93	8.90	19.50
Italy (C)	8.13	3.99	6.96	7.99	8.98	20.00
Netherlands (B)	14.30	7.72	11.20	13.40	16.60	42.10
Netherlands (C)	14.50	7.73	11.30	13.50	16.80	43.30

This table reports the estimation results from the panel part of the HFCS with only Belgium, Germany, Italy and Netherlands in the previous sample. The upper panel shows the cross validation results of the pooled results and the lower panel reports the country level results. The relative risk aversion is computed from equation $r_h = \frac{\mathbf{E}(r_m - r_f)}{\sigma_m^2} \frac{1}{G(X\hat{\theta})}$. The market risk premium is set to be 8%, and market volatility is set to be $(0.20)^2$. (B) stands for baseline model, and (C) stands for the cross validation results.

Table 8: Evidence from Survey of Consumer Finance 2016 in the U.S.

	Min	25%	Median	Mean	75%	Max
Estimate RRA						
All	2.272	3.990	5.810	6.847	9.587	12.377
Participants	2.272	3.073	3.663	3.934	4.762	5.800
Non-participants	2.347	5.319	7.425	8.778	12.560	15.515
Participation Costs $						
All	2.19	6.84	21.08	4563.37	354.54	245734.88
Financial Wealth 1000 $						
Participants	3.5	5.8	41.0	7797.2	902.5	1049273.5
Non-participants	3.5	4.5	10.4	2698.8	100.4	1047586.5
Total Wealth 1000 $						
Participants	-793.3	28.7	307.9	22529.6	4003.3	1641313.5
Non-participants	-381.4	6.6	113.7	10145.8	679.9	1527043.5
Comparison with previous methods:						
Friend and Blume 1975 method:						
	2.903	3.270	3.415	3.484	3.621	8.935
Chiappori and Paiella 2011 method:						
	3.188	3.865	4.142	4.243	4.515	10.128
Morin and Suarez 1983 method:						
	3.057	3.289	3.414	3.485	3.553	10.360

This table reports the estimation restuls of censored fractional response model from SCF 2016, and the comparison among the previous methods. The relative risk aversion is computed from equation $r_h = \frac{\mathbf{E}(r_m - r_f)}{\sigma_m^2} \frac{1}{G(X\hat{\theta})}$. The market risk premium is set to be 8%, and market volatility is set to be $(0.20)^2$.

Table 9: Results of Different Specifications – Part 1

	Baseline	Risky 1	Risky 2	$\tau = 0.02$	$\tau = 0.10$	Beta
Extremal Quantile Regression Coefficients:						
(Intercept)	-2.975	-3.026	-3.270	-3.288	-3.178	-2.975
Age	0.004	0.008	0.005	0.001**	0.006	0.004
Female	-0.110	-0.107	-0.111**	0.086**	-0.071**	-0.110
Education	0.103	0.091	0.171	0.102	0.149	0.103
Employement	-0.176	-0.183	-0.146	-0.074**	-0.168	-0.176
Income (log)	-0.137	-0.129	-0.144	-0.185	-0.072	-0.137
Censored Fractional Response Model Coefficients:						
Age	0.002	0.005	0.003	0.002	0.001	0.000
Female	-0.173	-0.147	-0.187	-0.162	-0.183	-0.173
Education	0.199	0.196	0.269	0.182	0.223	0.234
Marital	-0.217	-0.252	-0.246	-0.219	-0.213	-0.150
Family	-0.085	-0.079	-0.096	-0.080	-0.091	-0.085
Retiree	0.013**	0.044**	0.179**	-0.010**	0.045**	0.099
Employment	-0.258	-0.301	-0.168	-0.275	-0.235	-0.134
Wealth(log)	0.017	0.025	0.026	0.012	0.023	0.032
Income(log)	0.074	0.078	0.089	0.055	0.101	0.165
Residence	-0.380	-0.270	-0.201	-0.413	-0.336	-0.187
Housing Value	0.015	0.011	0.011	0.016	0.013	0.008
Other Prop	0.138	0.126	0.075	0.115	0.168	0.240
Country dummy	yes	yes	yes	yes	yes	yes

This table shows the estimation results of different empirical specifications. To make clean presentation, I signal the coefficients NOT significant at 5% level with **; all the other coefficients are significant. The column of "Risky 1" shows the result of treating bond as risky asset; the column of "Risky 2" shows the result of only considering stocks and mutual funds as risky assets; the column of "Beta" reports the result of using beta distribution for the errors.

Table 10: Results of Different Specifications – Part 2

	Baseline	Risky 1	Risky 2	$\tau = 0.02$	$\tau = 0.10$	Beta
Lower Censoring Thresholds – $L(\hat{X})$						
Min.	0.007	0.007	0.005	0.003	0.016	0.007
1st Quartile	0.015	0.016	0.013	0.007	0.029	0.015
Median	0.017	0.018	0.014	0.007	0.033	0.017
Mean	0.018	0.019	0.015	0.008	0.034	0.018
3rd Quartile	0.020	0.022	0.016	0.008	0.037	0.020
Max.	0.094	0.099	0.097	0.061	0.100	0.094
Participation Costs – δ_h^s						
Min.	0.0003	0.0003	0.0002	0.0001	0.0007	0.0003
1st Quartile	0.0006	0.0006	0.0005	0.0003	0.0012	0.0006
Median	0.0007	0.0007	0.0006	0.0003	0.0013	0.0006
Mean	0.0007	0.0008	0.0006	0.0003	0.0013	0.0007
3rd Quartile	0.0008	0.0009	0.0007	0.0003	0.0015	0.0007
Max.	0.0038	0.0040	0.0039	0.0024	0.0040	0.0033
Participation Costs in €						
Min.	0.0	0.0	0.0	0.0	0.0	0.0
1st Quartile	2.4	2.5	2.0	1.1	4.4	2.4
Median	10.5	11.1	8.7	4.5	19.9	10.5
Mean	69.7	76.8	60.7	26.4	148.0	69.7
3rd Quartile	38.7	41.5	33.0	16.1	76.0	38.7
Max.	28650	32530	25150	8430	76660	28650
Estimated Relative Risk Aversion:						
Min.	2.86	2.53	3.25	2.85	2.87	2.95
1st Quartile	6.74	4.89	7.74	6.46	7.12	7.77
Median	8.37	7.18	10.30	7.94	9.01	9.97
Mean	8.87	7.76	11.75	8.36	9.68	11.04
3rd Quartile	10.51	9.77	14.37	9.86	11.51	13.18
Max.	55.98	53.27	87.17	47.54	69.63	149.30

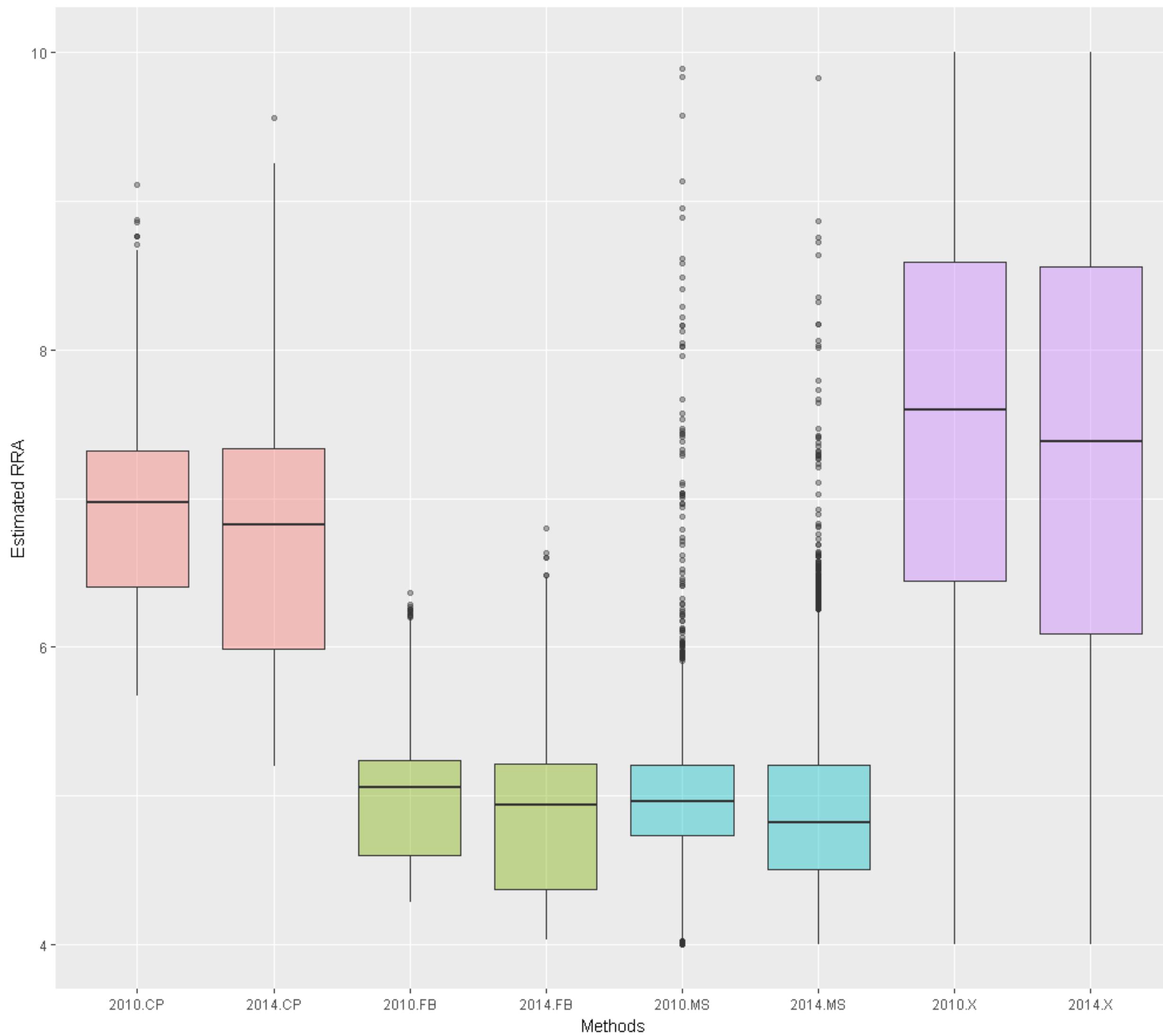
This table shows the estimation results of different empirical specifications. The column of “Risky 1” shows the result of treating bond as risky asset; the column of “Risky 2” shows the result of only considering stocks and mutual funds as risky assets; the column of “Beta” reports the result of using beta distribution for the errors.

Table 11: RRA with Different Perception of Risky Asset Markets: HFCS 2010

	Mean	Min	25%	Median	75%	Max
Expectation Differences:						
Optimistic Expectation: $\mathbf{E}[r_m - r_f] = 0.08$ and $\sigma_m = 0.20$						
	8.82	2.86	6.72	8.33	10.45	55.23
Pessimistic Expectation: $\mathbf{E}[r_m - r_f] = 0.04$ and $\sigma_m = 0.20$						
	4.41	1.43	3.36	4.16	5.22	27.6
Post War Average: $\mathbf{E}[r_m - r_f] = 0.07$ and $\sigma_m = 0.234$						
	5.62	1.83	4.29	5.31	6.66	34.99
Subjective Survey Expectation: $\mathbf{E}[r_m - r_f] = 0.06$ and $\sigma_m = 0.20$						
	6.59	2.14	5.03	6.23	7.81	41.06

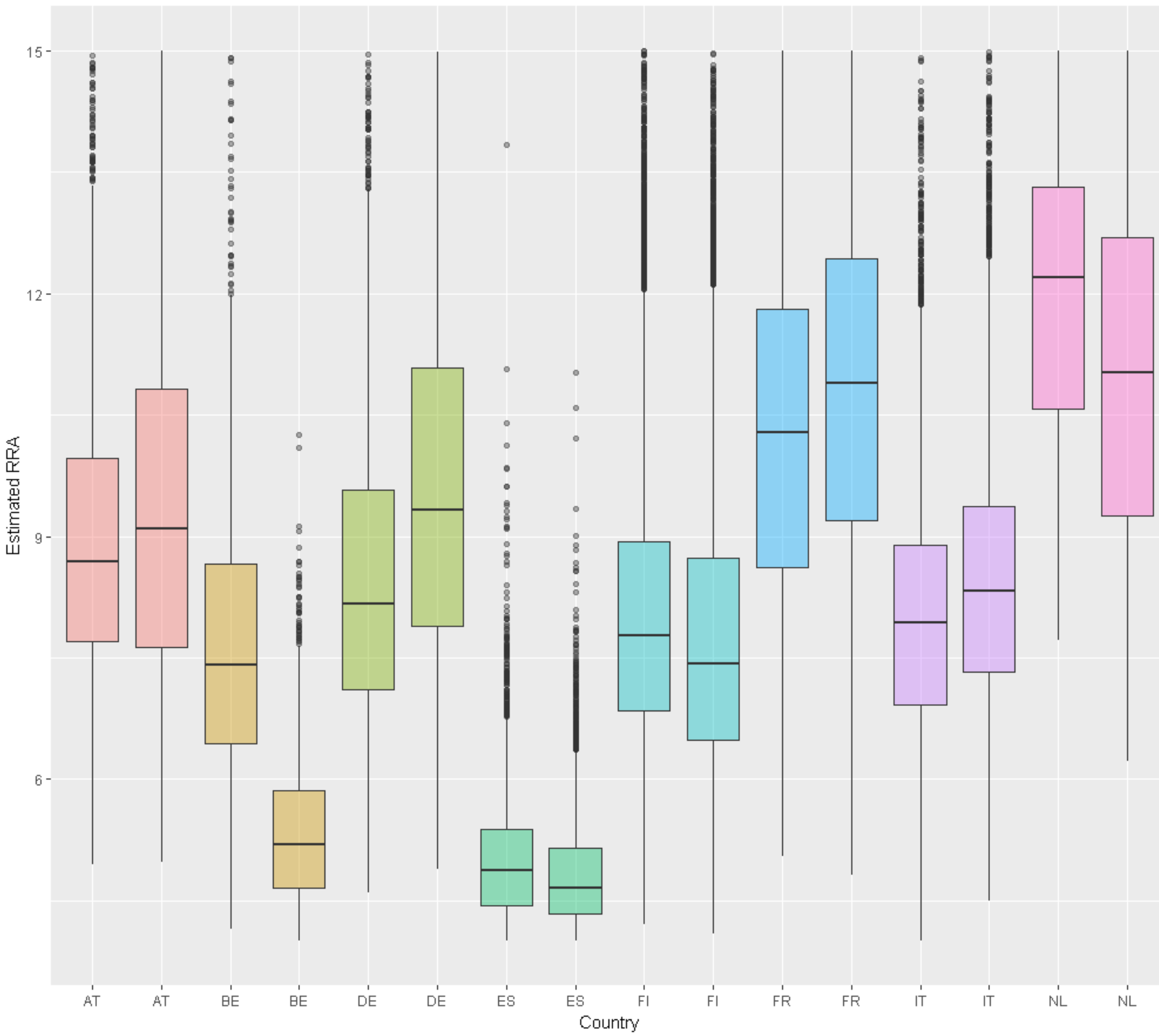
This table computes the estimated risk preference from HFCS 2010 using different market expectations. The relative risk aversion is computed from equation $r_h = \frac{\mathbf{E}(r_m - r_f)}{\sigma_m^2} \frac{1}{G(X\hat{\theta})}$. The market risk premium is set to be 8%, and market volatility is set to be $(0.20)^2$.

Figure 1: Relative Risk Aversion Estimates Comparison



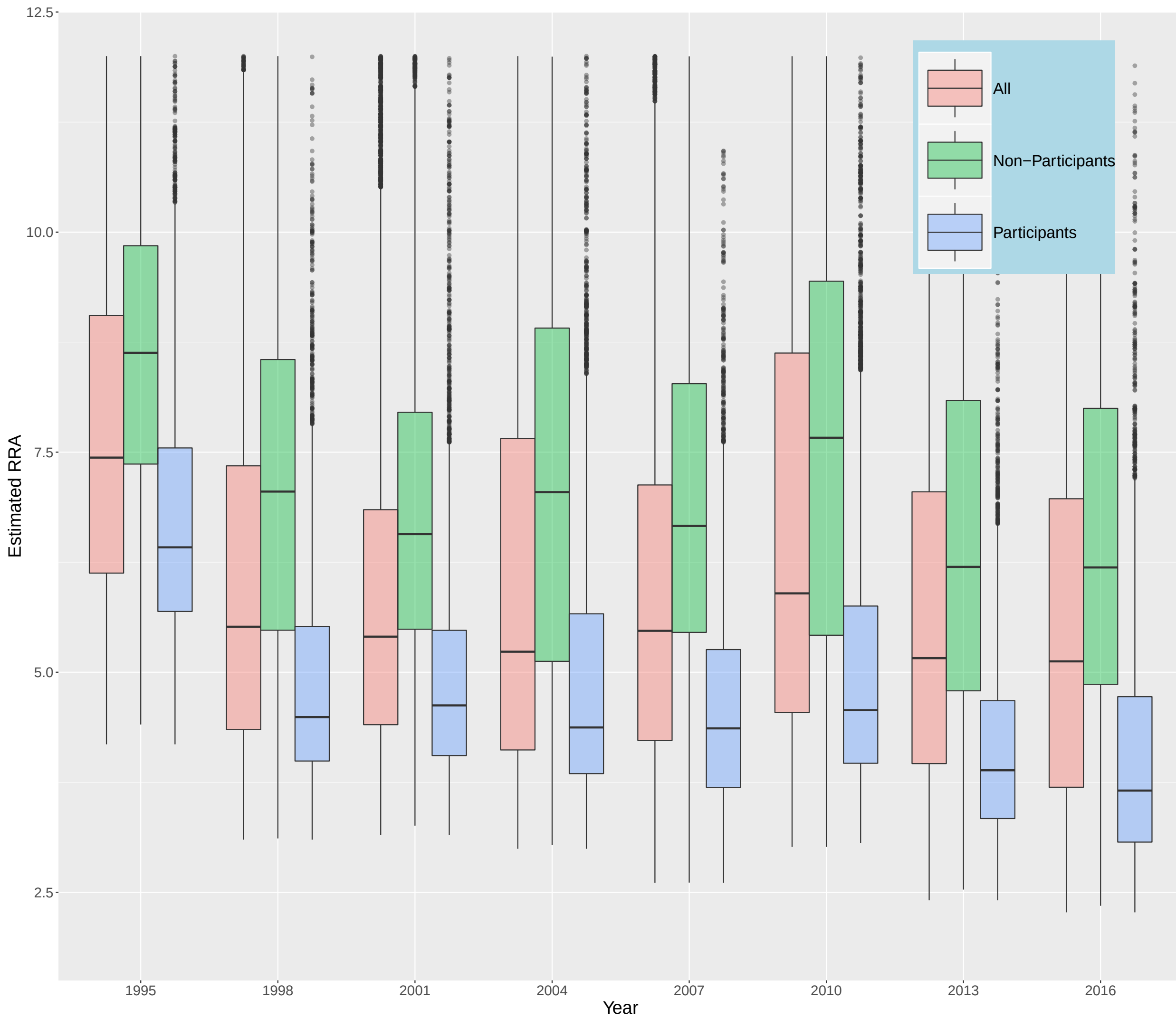
This figure depicts the difference in estimated relative risk aversion due to the methodogolies. X is the RRA estimates of this paper; FB is the RRA estimates of Friend and Blume (1975); MS is the RRA estimates of Morin and Suarez (1983); and CP is the RRA estimates of Chiappori and Paiella (2011). All methods apply to the two waves of HFCS. The maket price of risk $\mathbf{E}(r_m - r_f)/\sigma_m^2$ is 2 with risk premium $\mathbf{E}(r_m - r_f) = 0.08$ and maket volatility $\sigma_m^2 = 0.04$.

Figure 2: Relative Risk Aversion Estimates Country Comparison



This fiugre shows the estiamted relative risk aversion in eight European countries from the pooled estimation of the first two waves of HFCS. The lefthand side boxplot of each country comes from 2010; and the righthand side boxplot comes from 2014. The maket price of risk $\mathbf{E}(r_m - r_f)/\sigma_m^2$ is 2 with risk premium $\mathbf{E}(r_m - r_f) = 0.08$ and maket volatility $\sigma_m^2 = 0.04$.

Figure 3: Relative Risk Aversion Evolution from SCF 1995-2016



This fiugre shows the estiamted relative risk aversion from the Survey of Consumer Finance 1995–2016. In each wave, I report the distribution of the estimate relative risk aversion of all househoulds (red), non-participants (green) and participants (blue). The maket price of risk $\mathbf{E}(r_m - r_f)/\sigma_m^2$ is 2 with risk premium $\mathbf{E}(r_m - r_f) = 0.08$ and maket volatility $\sigma_m^2 = 0.04$.

A Choice between "complete portfolio" and "risky only"

Households with choice"risky only" invest all their financial wealth to risky assets. Their budget constraints are:

$$\mathbf{E}[W_{h1}] = W_{h0}\{1 + \mathbf{E}[r_m]\}.$$

By taking Taylor series expansion of $U(W_{h1})$ around W_{h0} and keeping the first two terms, the expected utility becomes:

$$\mathbf{E}[U(W_{h1})] = U(W_{h0}) + U^{'}(W_{h0})W_{h0}[\mathbf{E}[r_m] - (-\delta_h^b)] + \frac{1}{2}U^{''}(W_{h0})W_{h0}^2\sigma_m^2.$$

Compare with the optimal expected utility of "complete portfolio" in Equation 7:

$$V_b - V_0 = U^{'}(W_{h0})W_{h0}\{r_f + \alpha_h^*\mathbf{E}[r_m - r_f] - \mathbf{E}[r_m] - \delta_h^b\} + \frac{1}{2}U^{''}(W_{h0})^2(\alpha_h^{*2} - 1)\sigma_m^2.$$

Replacing α_h^{*2} with $-\alpha_h^* \cdot \frac{U^{'}(W_{h0})}{U^{''}(W_{h0})W_{h0}} \cdot \frac{\mathbf{E}[r_m - r_f]}{\sigma_m^2}$ from Equation 6 makes

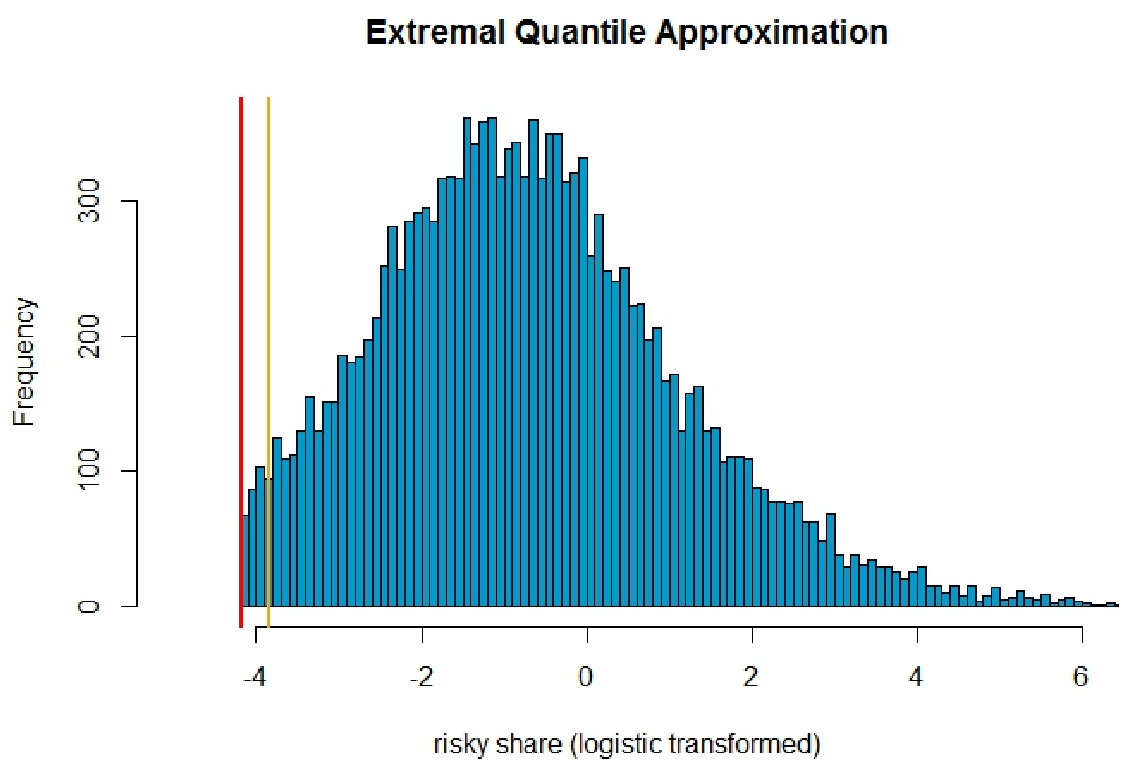
$$\begin{aligned} V_b - V_0 &= U^{'}(W_{h0})W_{h0}\{r_f + \alpha_h^*\mathbf{E}[r_m - r_f] - \mathbf{E}[r_m] - \delta_h^b - \frac{1}{2}\mathbf{E}[r_m - r_f]\frac{\alpha_h^{*2} - 1}{\alpha_h^*}\} \\ &= U^{'}(W_{h0})W_{h0}\{\mathbf{E}[r_m - r_f](\alpha_h^* - 1)(1 - \frac{1}{2}\frac{\alpha_h^* + 1}{\alpha_h^*}) - \delta_h^b\} \\ &= U^{'}(W_{h0})W_{h0}\{\frac{1}{2}\mathbf{E}[r_m - r_f](\frac{\alpha_h^*}{2} + \frac{1}{2\alpha_h^*} - 1) - \delta_h^b\} \qquad (18) \end{aligned}$$

Solve the equation $V_b - V_0 = 0$ and discard the solution outside of the interval $[0, 1]$, we would have the upper censoring function $H(X) \equiv (\frac{\delta_h^b}{\mathbf{E}[r_m - r_f]} + 1) - \sqrt{(\frac{\delta_h^b}{\mathbf{E}[r_m - r_f]})^2 + \frac{2\delta_h^b}{\mathbf{E}[r_m - r_f]}}$.

B Intuition of using extremal quantile approximation

The intuition of using extremal quantiles to approximate the censoring thresholds is as follows. According to the participation condition $\frac{1}{2}\mathbf{E}[r_m - r_f]\alpha_h^* - \delta_h^s \geq 0$, when participation cost is the same for every household, there exists a cut off threshold of risky shares below which should have no observation. Then the lowest positive observation is the estimate of such a cut off thresholds. However, this is not exactly the true cut off threshold. It is only certain that the true cut off threshold is below the estimate. Consider the sparse observations at the tail; it is not likely that we can have a precise approximation of the threshold with the limited sample size. Thus, I use the extremal quantile estimation to have a more precise approximation that is sufficiently close to the cut off threshold. Since the extremal quantile regression uses a simulation-based estimation using the sub-sample re-sampling technique to have the unbiased estimation of the extreme quantiles. Moreover, when the sample size is sufficiently large, we can push the extreme quantile to zero, which is exactly the minimum. Figure B.1 demonstrates the intuition of approximation of extreme quantile. I assume that with enough controls in the extremal quantile regression, the unobserved heterogeneity can be ignored.

Figure B.1: Intuition of Extremal Quantile Approximation



This method circumvents the difficulty of directly estimating the unobservable censoring thresholds. Instead, it approximates the thresholds with the extreme quantiles at both ends. However, with the normality assumption for the baseline specification, extremal quantile regression is not the only approach that we can obtain the estimation of the model. As long as we have a non-linear functional form of the cumulative distribution function, we can estimate all the parameters of the model using maximum likelihood. Then this approach is close to a control function approach in a simple sample selection problem. This paper insists on using extremal quantile regression for its flexibility and applicability in different empirical specifications.

Another issue that justifies the approximation of extreme quantile. Although Household Finance and Consumption Survey is well administrated, there are still a significant amount of confusing observations, such as the ones holding 20 euros in all types of risky financial assets when one's total financial wealth is well above 5000 euros. Those observations cannot be deemed rational simply consider the existence of transaction fee. In this paper, households who have less than 100 euros in risky assets are considered to be measurement errors. Therefore the minimum of risky shares is partly arbitrary.

C Additional Summary Statistics

Households with different demographic or personal status make different financial decisions as seen in Table C.2 and Table C.3. First of all, the total net wealth is one of the most important and interesting elements in the discussion about risky attitudes, starting from the early works of Morin and Suarez (1983) and Cohn, Lewellen, Lease, and Schlarbaum (1975) to the more recent works by Brunnermeier and Nagel (2008) and Chiappori and Paiella (2011). The relationship between wealth and risk aversion offers insights into the true risk preference with respect to wealth. More recently, studies using panel data note that a constant relative risk aversion may be the right description as they find that the change of wealth does not seem to affect the risk attitude. However, cross sectional study often finds the opposite effect. From the first part of Table C.2 and Table C.3, it is evident that rich households are more active in the financial market, with higher participation rates and higher shares in risky assets such as mutual funds and stocks, especially in the top 20% income households. Moreover, rich households also have higher shares of their total financial wealth into risky assets such as mutual funds, bonds, and stocks. This may imply decreasing relative risk aversion as well as more guided and sophisticated investing. Such a pattern is consistent with most of the empirical findings in the literature of household finance.

The life-cycle of portfolio choices is another important issue in household finance. People have a finite horizon in working and biological lives and adjust their portfolios as they age. Empirical evidence on the life-cycle of portfolio choice can be found in applied micro-econometric studies and macroeconomic calibration works such as Cocco et al. (2005), Fagereng et al. (2013) and Gomes and Michaelides (2005). In the second part of the tables, we can see that participation in all types of financial markets steadily increases until retirement and starts to fall afterward. Such a finding consistent with both previous findings and the theoretical explanation that households start to deplete their financial wealth when they no longer have labor income. As people age, they tend to invest relatively more shares of their financial wealth in risky assets and much less in pension and life insurance. The composition of financial assets tilts towards mutual funds, bonds, and stocks. It is easy to understand that after retirement, the sources of income are mainly in financial assets and pensions. This may be why retired individuals have higher shares of investment in risky assets than working individuals.

Housing is a major background risk, as discussed in Heaton and Lucas (2000); this has a significant impact on household portfolio choices. Homeowners, with or without a mortgage, are more confident and more willing to invest in financial assets. Among homeowners, households with a housing mortgage tend to hold fewer shares in risky assets but more in pensions and behave more like renters. However, this may also be attributed to the fact that outright homeowners are usually much wealthier than those with standing mortgages. Employment status makes a difference in the participation rate of financial assets in that both employees and self-employed individuals have significantly higher participation rates compared to retirees and unemployed individuals. However, employment does not make a significant difference in portfolio choices. A temporary employment shock should not affect a person's general risk preference. Finally, education plays a major role in financial behavior, especially regarding the investment in mutual funds and stocks. Participation rate doubles at each level of education. People with a college education are four times more likely to participate the mutual funds and stocks. Higher education also leads households to invest more shares of their total financial wealth to risky assets.

Apart from the demographic differences, the first wave of HFCS also shows significant country differences. Table C.1 presents the country profiles on the seven financial categories, which consists the median real values[10], the shares of financial assets and participation rate to financial assets. For the total financial portfolio sizes, those selected countries differ greatly. The Netherlands and Belgium invest more in the financial market, while households in Spain, Finland and France do not invest as much. The difference in income cannot account for all the variance of the amount of investment across countries as most of the selected countries have similar GDP per capita. Households in different countries have different preferences across financial assets as well. Bonds are considered to be unattractive in most countries, and this is reflected in both participation rates and investment shares. Interestingly, bonds are quite popular in Belgium and Italy. As for private pension and whole life insurance, households in France, Germany, and Netherlands are much more enthusiastic than in other countries. Finland's stock market participation stands out from the rest with a participation rate as high as 22.2%. Different investment behaviors at the country level imply that there might be country specific features affecting households' investment decisions, such as the cultural difference, the financial institution or legal system difference. For instance, countries with a better protection of renters usually have lower ownership of residence.

[10]Note that the median value for all the categories does not add up to the total. That is because the conditional median value is computed separately in each category. In other words, conditional on participation means participation to that very financial asset, not all the financial assets.

D Estimation with Beta Distribution

The alternative method of parametrize the censoring probability is to assume that the distribution of the unobserved heterogeneity in risky shares is a Beta distribution. Therefore the censoring probabilities P_M, P_L and P_H take the functional form as follows:

$$\begin{aligned} P_L(X,Z) &= Pr\{G(X_h\theta_M) + u_h \leq G(Z_h\hat{\theta}_L)|\hat{\theta}\} = I_{G(Z_h\hat{\theta}_L)}(\alpha,\beta) = \frac{\mathbf{B}(G(Z_h\hat{\theta}_L),\alpha,\beta)}{\mathbf{B}(\alpha,\beta)} \\ P_H(X,Z) &= Pr\{G(X_h\theta_M + u_h) \geq G(Z_h\hat{\theta}_H)|\hat{\theta}\} = I_{G(Z_h\hat{\theta}_H)}(\alpha,\beta) = \frac{\mathbf{B}(G(Z_H\hat{\theta}_L),\alpha,\beta)}{\mathbf{B}(\alpha,\beta)} \\ P_M(X,Z) &= 1 - P_L(X,Z) - P_H(X,Z) \end{aligned}$$

where $I_x(\alpha,\beta)$ is the regularized incomplete beta function, which is the cumulative distribution function of beta distribution, $\mathbf{B}(x,\alpha,\beta)$ is the incomplete beta function, and $\mathbf{B}(\alpha,\beta)$ is the beta function.

Following the suggestion of Ospina and Ferrari (2012), I re-parametrize the beta distribution with μ and ϕ ($\mu \in (0,1)$ and $\phi > 0$). The relation with the original parameters α and β are as follow:

$$\begin{aligned} \mu &= \tfrac{\alpha}{\alpha+\beta} \\ \frac{\mu(1-\mu)}{(\phi+1)} &= \tfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \end{aligned}$$

Then for the beta distribution, μ is the distribution mean and $\frac{\mu(1-\mu)}{(\phi+1)}$ is the variance, in which ϕ plays the role of a precision parameter. With such parametrization, by making $\mu = G(X_h\theta_M)$, we can compute the censoring probability easily. ϕ will be estimated along with the rest of the parameters in the censored fractional response model.

Table C.1: Household Finance Summary by Country

	Total	Deposits	Mutual Funds	Bond	Stocks	Money Owed	Pension & Life	Other
Austria(2010)								
Real value	13.5	10.6	11.2	13.8	7.1	2.6	8.1	7.7
Shares	100.0	63.5	11.8	6.9	3.1	3.5	8.9	2.2
Participation	99.5	99.4	10.0	3.5	5.3	10.3	17.7	1.6
Belgium(2010)								
Real value[1]	26.5	10.0	20.4	30.8	5.1	2.3	19.9	21.0
Shares[2]	100.0	39.1	13.0	14.8	10.4	1.5	16.7	4.5
Participation[3]	98.0	97.7	17.6	7.5	14.7	7.7	43.3	3.5
Germany(2010)								
Real value	17.1	7.9	10.0	16.0	8.6	2.7	11.4	2.1
Shares	100.0	44.4	10.4	5.6	6.5	2.7	26.8	3.6
Participation	99.3	99.0	16.9	5.2	10.6	13.7	46.5	11.3
Spain(2008)								
Real value	6.0	3.5	13.9	19.2	6.1	6.0	7.4	12.0
Shares	100.0	51.4	7.7	1.9	9.1	6.4	15.1	8.4
Participation	98.3	98.1	5.6	1.4	10.4	6.3	23.6	1.9
France(2010)								
Real value	10.7	6.5	6.9	12.0	6.9	3.0	10.6	5.0
Shares	100.0	33.8	5.8	1.4	11.6	1.0	39.0	7.4
Participation	99.6	99.6	10.7	1.7	14.7	5.0	37.5	7.8
Finland(2009)								
Real value	7.4	4.5	2.6	10.0	3.8	NA	4.3	NA
Shares	100.0	51.9	11.5	1.0	26.1	NA	9.5	NA
Participation	100.0	100.0	27.4	0.8	22.2	NA	23.7	NA
Italy(2010)								
Real value	10.0	5.9	20.0	20.0	10.9	4.0	10.1	10.4
Shares	100.0	46.9	9.6	20.4	4.5	0.5	8.8	9.3
Participation	92.0	91.8	6.3	14.6	4.6	1.3	18.0	3.7
Netherlands(2009)								
Real value	34.7	10.1	7.1	15.5	5.6	2.0	53.2	5.5
Shares	100.0	33.9	6.4	4.3	3.5	1.7	49.3	0.9
Participation	97.8	94.2	17.7	6.0	10.4	8.5	49.8	2.7

[1] Median value in thousand EUR conditional on participation.
[2] Average shares of all type financial asset conditional on participation.
[3] Percentage of households participating specific type of financial assets.

Table C.2: Participation in Financial Assets by Household Demographics

	Total	Deposits	Mutual Funds	Bond	Stocks	Money Owed	Pension &Life	Other
Euro Zone	96.8	96.4	11.4	5.3	10.1	7.6	33.0	6.0
Percentile of Net Wealth								
less than 20	93.2	92.5	2.0	0.2	1.2	7.8	15.9	1.7
20-39	96.7	96.3	8.1	1.7	5.0	10.2	32.7	4.6
40-59	96.4	96.1	10.4	3.9	8.0	5.9	31.5	4.7
60-79	98.4	98.1	12.4	6.6	11.0	5.7	35.8	5.4
80-100	99.5	99.1	23.8	14.0	25.2	8.6	49.1	13.8
Age of Reference Respondent								
16-34	97.4	97.1	9.7	1.7	6.7	10.3	33.7	4.8
35-44	97.5	97.0	12.9	3.4	10.1	9.0	41.1	6.3
45-54	97.0	96.7	13.0	5.0	11.2	8.0	43.7	5.4
55-64	97.2	96.4	13.1	7.6	13.3	7.5	37.7	7.4
65-75	96.4	96.1	10.9	8.1	10.4	5.8	19.4	7.3
75+	95.0	94.7	6.9	6.6	7.6	4.2	12.8	4.9
Work Status of Reference Respondent								
Employee	97.9	97.6	13.3	4.2	11.4	7.9	42.3	5.7
Self-Employed	96.9	96.6	12.7	7.9	12.5	12.6	44.7	10.4
Retired	95.9	95.6	9.4	7.5	9.3	5.5	19.0	6.4
Other No Work	94.9	94.1	6.8	1.5	3.8	8.6	21.9	3.0
Housing Status								
Owner-Outright	96.6	96.3	11.9	8.9	12.4	5.1	28.9	6.3
Owner-Mortgage	98.7	98.1	16.2	3.7	13.6	7.8	47.8	7.4
Renter or Other	96.2	95.7	8.5	2.4	6.0	10.1	30.1	5.2
Education of Reference Respondent								
Primary or Non	93.6	93.1	4.0	4.0	4.2	4.5	19.0	2.4
Secondary	98.2	97.9	10.8	5.2	9.2	8.9	36.4	6.1
Tertiary	99.0	98.7	22.6	7.2	19.6	9.9	46.8	11.1

Table C.3: Shares in Financial Assets by Household Demographics

	Total	Deposits	Mutual Funds	Bond	Stocks	Money Owed	Pension &Life	Other
Euro Zone	100.0	42.9	8.7	6.6	7.9	2.2	26.3	5.3
Percentile of Net Wealth								
less than 20	100.0	65.7	1.8	NA	1.2	4.4	26.1	0.6
20-39	100.0	62.3	5.4	1.4	1.7	3.9	23.9	1.3
40-59	100.0	55.4	5.5	2.5	2.9	1.9	30.1	1.7
60-79	100.0	53.5	6.7	4.0	4.1	1.8	28.2	1.7
80-100	100.0	35.4	10.4	8.6	10.6	2.2	25.4	7.4
Age of Reference Respondent								
16-34	100.0	56.6	5.1	1.1	4.6	1.7	26.3	4.3
35-44	100.0	43.3	6.8	3.5	7.0	2.9	30.0	6.4
45-54	100.0	40.4	8.8	3.9	6.7	2.8	32.7	4.7
55-64	100.0	39.0	9.9	7.1	7.7	2.0	27.9	6.3
65-75	100.0	44.0	10.7	10.0	10.4	2.2	18.3	4.4
75+	100.0	46.0	7.6	10.6	9.4	1.3	20.2	4.8
Work Status of Reference Respondent								
Employee	100.0	44.4	8.2	3.8	7.1	1.7	30.3	4.4
Self-Employed	100.0	34.0	8.3	6.6	8.8	3.8	27.4	11.2
Retired	100.0	45.2	9.4	9.8	9.0	2.0	20.5	4.2
Other No Work	100.0	46.4	11.0	4.3	4.9	3.5	27.6	2.4
Housing Status								
Owner-Outright	100.0	43.5	8.7	8.6	9.1	1.7	22.4	6.0
Owner-Mortgage	100.0	40.3	7.8	2.7	6.4	2.9	35.9	4.0
Renter or Other	100.0	43.8	9.7	4.6	6.3	3.1	27.8	4.7
Education of Reference Respondent								
Primary or Non	100.0	51.3	5.1	7.1	4.7	2.5	26.1	3.1
Secondary	100.0	45.6	7.1	6.3	6.6	2.0	27.9	4.5
Tertiary	100.0	37.7	11.4	6.5	10.1	2.3	25.2	6.7