Departamento de Economía
Facultad de Ciencias Sociales
Universidad de la República

Documentos de Trabajo

Estimating heterogeneous costs of participation in the risky asset markets

Graciela Sanroman

Documento No. 21/07
Noviembre, 2007

Estimating heterogeneous costs of participation in the risky asset markets

Graciela Sanroman*
CEMFI and UDELAR
PRELIMINARY AND INCOMPLETE

First draft: Junio, 2007

Abstract

This paper develops and estimates a dynamic structural model of participation in the risky financial asset markets using household level panel data. We specify a simple economic model in order to capture the portfolio choice over the life cycle. We solve the model using numerical techniques. Then we embed the optimal solution into the statistical (auxiliary) model and estimate the structural parameters using Generalized Indirect Inference. This paper focuses on the estimation of the non proportional costs to participate in the risky asset markets. We consider heterogeneous costs among education groups. We find that participation costs in the risky asset markets are positive and significant. We also conclude that they vary a lot among education groups.

JEL classification: C15, C61, D14, D91, G11

Keywords: Portfolio choice, dynamic programming, indirect inference.

*Thanks are due to Manuel Arellano for numerous insights and helpful suggestions. Also, I thank Victor Aguirregabiria, Pedro Mira and Josep Pijoan-Mas for their useful comments.

1 Introduction

The literature on household portfolio choice is gaining richness and complexity. Two main conclusions arise from the recent research on this issue. First, there is not a simple rule to select an optimal financial portfolio over the life cycle. Second, actual household portfolio behaviour is extremely heterogeneous and most of this heterogeneity is given by the decision of whether to participate in the risky asset markets.

This paper develops and estimates a dynamic structural model of participation in the risky asset markets using household level panel data. We follow Deaton (1991) consumption model and extend it by introducing the decision on the risky asset demand. To this end we implement a simulated based estimation method which requires solving the economic problem of each individual in the sample. There is not an available analytical solution to this problem, thus, our solution is based on numerical techniques.

In recent years some authors have solved numerically life cycle models of consumption and portfolio (see e.g. Viceira, 2001 and Cocco *et al*, 2005). We take a different route from them by estimating our model using microdata. The method we use is the Generalized Indirect Inference (Keane and Smith, 2003). To our knowledge this is the first time that such a formal econometric approach is used in the field of household's portfolio choice.

The data come from the Italian "Survey of Households Income and Wealth". This is a unique dataset which cointains panel data on household wealth, income and consumption and also includes detailed information about financial holdings and the demographic characteristics of households members. Moreover, its quality have been proved to be good enough for our purposes.

The existing literature shows a variety evidence about households financial decisions (Campbell, 2006; Campbell and Viceira, 2002 and Guiso and Japelli, 2001): (i) households are extraordinarily diverse in their portfolio choices, (ii) lack of participation, i.e., the zero solution in the demand for the risky assets, accounts for most of this heterogeneity, (iii) the average real return on stocks and bonds has been higher than that on bank deposits, (iv) wealthy people are more likely to hold stocks but a substantial proportion of wealthy people hold no stocks at all, (v) those who hold risky financial assets were more likely to have these types of holding in the past, (vi), the participation rate varies considerably with age but, (vii) age effects on portfolio choice are very difficult to identify because time and cohort effects are also present, and finally, (viii) many households do not diversify their risky asset holdings. We focus on the analysis of the first six aspect above but we ignore the last two by assuming that there are only two assets in the economy and neither time nor cohort effects are present.

Viceira (2001) and Jagannathan and Kocherlakota (1999) argue that the common advise of some financial planners of reducing the risky asset share as ageing is supported by the theory. On the contrary, our results show that the average of the optimal risky asset shares does not depend on the age of the agent.

Viceira (2001) and Cocco *et al* (2005) conclude that the optimal risky asset share increases in the presence of labor income and it is a decreasing function of current wealth for a given future labor income stream. We find some evidence to support these conclusions.

Nevertheless, those authors do not consider the corner solution at zero and actually most of the observable heterogeneity relies on zero versus non-zero solutions. Perraudin and Sorensen (2001), Halliasos and Michaelidis (2001), Miniaci and Weber (2001), Vissing-Jorgensen (2002) Paiella (2006) and Attanasio and Paiella (2006) propose to introduce non proportional participation costs into financial risky asset demand in order to explain zero solutions. We find that zero solution is an optimal choice only if both borrowing restrictions and participation costs are present. Moreover, agents' preferences and beliefs about both future non-financial income and future returns on financial investments play a crucial role in their demand for the risky assets.

Briefly, our model considers a rational, risk-averse agent who derives utility from consumption. The agent is endowed with human capital and financial wealth and she has to decide how much to consume at the present and where to invest her savings. In order to select the optimal alternative the agent brings the future into the picture and builds contingent plans of consumption and investment for her remaining lifetime.

The model could help us to exploit the information contained in the dataset with the aim of learning about agents' financial behaviour, their attitude towards risk and the effect of the existence of uncertainty, borrowing restrictions and participation costs. It will also enable us to analyze how some of the agent's observable characteristics (wealth, non-financial income, education and age) affect her demand for the risky asset. However, in this paper we focus almost exclusively on the estimation of non proportional participation costs to invest in the risky asset markets. We proceed by assuming that households choose the optimal portfolio and calculate how much cost is needed in order to explain the observed households' behaviour. We consider heterogeneous costs among education groups. The education group is determined by the education level of the household's head. We distinguish four groups: elementary school, secondary school, high school and college or post-graduated.

We found that participation costs in the risky asset markets are positive and significant. We also conclude that they vary a lot depending on the education of the household's head. Our results are comparable with those Paiella (2006) and Attanasio and Paiella (2006) found using US data.

The rest of the paper is organized as follows. Section 2 presents the data and analyzes the main facts in the issue of the Italian households' financial portfolios. In Section 3 the economic model and the solution method are described. In Section 4 the econometric approach is formally established. Section 5 presents the estimation of the non-financial income process and establishes the assumptions about the asset returns and the preference parameters. Section 6 includes the results of the structural estimation. Section 7 concludes.

2 Data and main facts

Datasets including household wealth information are rare. However, the number of countries that are collecting this type of information is increasing quickly during recent years. There are some available data on household financial wealth from Canada, Germany, Japan, Spain, UK, US, the Netherlands but we have concentrated our search on the Italian data.

We use the Italian survey called "Survey of Households' Income and Wealth" (SHIW). It has been collected by the Bank of Italy since 1965, but it is only since 1987 that it includes the minimum set of information we need for our purposes here. At first it was conducted annually but since 1987 it has been following a biannual frequency. Nevertheless, the wave of 1998 took place three years later than the previous one. The latest available wave corresponds to 2004. In the 1989 wave the 15 percent of the 1987 respondents was reinterviewed and since this year a rotating panel structure has been maintained. Now we can obtain up to 9 observations for some households, but only 10 of such households are available. However, the sample size increases a lot (it is greater than 3000) if we look for a panel with at least 4 observations per household.

The SHIW collects information on family income, savings, asset holdings, durable goods (expenditures and holdings), other real assets. It also includes detailed information on the composition of Italian households' wealth, both real and financial. We classify financial asset categories among three sets (i) deposits, (ii) bonds and stocks (iii) private ownership and focus the analysis on the first two of these categories.

Table 1 show some descriptive statistics for the whole sample. The respective information on the used sample is included in Table 2. It is important to point out that the selection of the used sample is due to the availability of information. We need households with at least four observations and with not missing data in the relevant variables of the model.

Some conclusions arise from the analysis of Tables 1 and Figure 1. First, the proportion of risky asset holders has increased substantially since 1987 to 1995 (from 24 to 40 percent) but then it decreased. The picture in 2004 was similar than that of the beginning of the nineties: only one of three households own bonds and/or stocks.

In Table 2 and Figure 1 we observed that the used sample differs from the whole sample. In particular, the percentage of risky asset holders is higher and it evolution does not show a hump shaped behaviour. The proportion of participants attains to 36 percent in 1993, rose to 47 percent in 1995 and then fluctuated. In 2004 it was 40 percent.

There are also differences between descriptive statistics in the whole sample and those in the used one. The average of the age of the household head's is 52.6 in the whole sample and 57.1 in the used one. The averages of non-financial income, financial holdings and real wealth are 25230, 19841 and 126781 euros (1995 prices) respectively in the whole sample and 27611, 26890 and 168939 in the used one.[1]

Figures 2 and 3 illustrate the age-portfolio profile in the used sample. Figure 2 shows that the risky asset shares are very heterogeneous and that the most important difference is due to the proportion of risky asset holders. In Figure 3 we identify a weak humped-shaped profile for the average proportion of participants over the life cycle.

Figures 4 and 5 illustrate the portfolio profile against the cash-on-hand to non-financial income ratio (in logs). Figure 4 shows that both zero and one corner solutions are present in the data. We also observe a lot of het-

[1]Non-financial income includes all sources of income except for financial assets. It includes compensation of employees, pensions an other transfers, net income from self-employment and entrepreneurial income and income from buildings including actual and imputed rents.

erogeneity among household portfolio choices. Instead, in Figure 5 a clear increasing pattern for the proportion of participants by deciles of the cash-on-hand/non-financial income ratio is observed. However, a substantial portion (45 percent) of the wealthy people does not participate in the risky asset market.

Figures 6 and 7 illustrate the non-financial income-portfolio profile. In Figure 6 we find heterogeneity but it is clear that the risky asset share of those in the top of the non-financial income distribution is greater than that of those in the bottom. Analyzing the behaviour among deciles of non-financial income, in Figure 7, we also observe a clearly increasing pattern.

Figures 8 and 9 analyze the risky asset shares by education group. An increasing pattern is also clear in these graphs. The increasing pattern is remarkable: solely less than 20 percent of the household in the bottom of the education distribution participates in the risky asset markets while the proportion of the better educated rises to 50 percent.

Finally, Table 3 reproduces the results of Pelizzon and Weber (2007). These authors document that in Italy during the period 1989-2003 expected returns of Government Bonds, Corporate Bonds and Stocks attain 4.1, 2.3, and 4.9 respectively. That implies that in Italy there was a premium on the returns of the financial risky asset in the recent years.

Summarizing, the evidence on Italians household portfolio are in line with the main facts indicating by the existing literature (Guiso and Japelli, 2001). Households are extraordinarily diverse in their portfolio choices and lack of participation accounts for most of this heterogeneity. Wealthy people are more likely to hold stocks but many wealthy people do hold neither stocks nor bonds. Furthermore, the participation rate varies considerably with the non-financial income and cash-on-hand/non-financial income ratio. The well educated tends to participate more in the risky asset markets. However, we find no evidence of any relationship between the risky asset share and the age of the agent.

3 Economic model

3.1 Model specification

The model we specify is the simplest one we can imagine in order to capture the portfolio choice over the life cycle when there are other sources of income (wages, pensions, real state rents) apart from financial assets returns. It considers a risk-averse agent who derives utility from consumption, that is to say, her goal is to maximize expected discounted utility over her remaining lifetime. There are only two type of assets in the economy: human capital and financial assets. The agent's life horizon is finite and there are no bequest motives. This model has some shortcomings, for example the effects of housing (Cocco, 2004 and Pelizzon and Weber, 2007); entrepreneurial risk (Heaton and Lucas, 2000) and retirement (Viceira, 2001) are either ignored or treated in a stylized manner.

We assume that at the beginning of period t the agent receives a liquid endowment. That endowment consists of financial wealth (W_t) and non-financial income (nfi_t). Financial wealth comes from the realized gross return on the investment made in the previous period. Non-financial income comes from a deterministic process which is determined by the agent's human capital. We assume that both human capital and the respective working supply are unobservable and exogenously given. Consequently the non-financial income is also exogenous. In spite of the fact that there is no moral hazard in that context, we assume there are no insurance markets for non-financial income, so the agent is not allowed to borrow.

In each period t, the agent simultaneously decides (i) how to allocate resources among consumption and savings; and (ii) the demand for the risky asset as a fraction of her financial holdings. There are two different assets in the economy. Both the expected return and the volatility on the asset that we have called "the risky asset" are higher than those on the another, which we named "the risk-less asset". Participation in the risk-less asset market is completely free of cost but this is not so in the risky asset market. We assume there are some non proportional costs associated with the participation in the risky asset market and that participation cost is lower for those agents who participated in that market during the previous period.

The problem can be briefly described as follows. In each time period t the agent observes her income sources, the composition of her portfolio in the previous period and the set of investment opportunities available in the market. Then, based on her preferences and beliefs about future income and asset returns and taking into account that she is not permitted to borrow, the agent chooses her consumption expenditure (which also determines financial holdings) and the proportion of financial holdings to allocate to the risky asset. In order to select the optimal alternative the agent brings the future into the picture and builds contingent plans of consumption and investment for her remaining lifetime. We assume the agent's preferences are described by a standard, time separable, utility function over consumption ($u(c)$).

Thus, the agent's problem in period t is to choose c_t, the proportion of the cash on hand to spend in consumption (alternatively A_t the wealth transferred to the next period) and α_t, the portion of the financial wealth to be allocated to the risky asset.

3.1.1 Preferences

Time is discrete, each period corresponds to two years. The agent lives until age H with certainty. Utility is time separable and the utility function is one of the power utility function.

$$V_t = \sum_{s=t}^{H} B^{s-t} u(c_s) \tag{1}$$

$$u(c_s) = \begin{cases} \frac{c_s^{1-crra}}{1-crra} & \text{if } s \leqslant H \\ 0 & \text{if } s > H \end{cases} \tag{2}$$

where $crra$ is the coefficient of relative risk aversion, B is the discount factor and c_s is the consumption at age s. We assume there is not bequest motive and thus set the value of the utility function to zero for those years where age is greater than H.

3.1.2 Nonfinancial income process

We assume the future non-financial income stream is deterministic and that the agent knows it with certainty. However, the econometrician could only observe the non-financial income at time s with $s < t$. Thus from the point of view of the econometrician the nonfinancial income process is random. We assume this stochastic process is given by,

$$nfi_{is} = x'_{1is}\pi_1^{inc} + f_i\pi_2^{inc} + t\pi_3^{inc} + h_{is} \quad s = t, t+1, ...H \tag{3}$$

where

nfi_{is} : nonfinancial income of the household

x'_{1is} : age and age square of the household' head

η_i : an unobservable time invariant individual characteristic

$f_i^{'}$:dummies of household's head education level, a dummy indicating if the household resides in the South of Italy and dummies indicating the city size.

t a polynomial in time and a dummy to 1998.

h_{is} is a stochastic term that is completely defined in Section 5

3.1.3 Financial assets

We assume that there are two asset in which the agent can invest, a risk-less and a risky asset. The investment opportunity set is constant over time. The risk-less asset (cash and deposits) has a constant gross return R^f. The risky asset (stocks and bonds) return is denoted by R_s. This return is random and there is an expected excess return $E_t\left(R_{t+1} - R^f\right) > 0$.

3.1.4 Participation costs

The investment in the risky asset is costly. There are both monitoring and entry cost that the agent must pay, and there are two type of costs: monetary and opportunity. These costs are non proportional to the amount invested in the risky asset. The cost function is denoted by $G_s = G(\alpha_s, Z_s)$ where α_s and α_{s-1} are the risky asset shares at periods s and $s+1$ and Z_s are some characteristics of the agent. We assumet that G_s equals zero if $\alpha_s = 0$.

3.1.5 Wealth accumulation

The financial wealth at the beginning of each period is denoted by W_s and equals savings in the previous period times the gross return of the financial portfolio.

$$
\begin{aligned}
W_s &= A_{s-1} \times R(\alpha)_s && (4)\\
A_s &= coh_s - G_s - c_s && (5)\\
coh_s &= W_s + nfi_s && (6)\\
R(\alpha)_s &= R^f + \alpha_{s-1}(R_s - R^f) && (7)\\
G_s &= \begin{cases} 0 & if\ \alpha_s = 0 \\ G_s^*(Z_s) & if\ \alpha_s > 0 \end{cases} && (8)
\end{aligned}
$$

Following Deaton (1991) household can not borrow against future labor income.

$$A_s \geq 0 \tag{9}$$

Moreover, it is not possible to borrow in the risk-less asset market with the goal of investing in the risky asset market.

$$0 \leqslant \alpha_t \leqslant 1 \tag{10}$$

3.2 The optimization problem

The optimization problem can be written as

$$\max_{\{c_s,\alpha_s\}} E_t \left\{ \sum_{s=t}^{H} B^{s-t} u(c_s) \right\} \tag{11}$$

subject to the constraints given by equations [4] to [10].

To our knowledge, there is not any available analytical solution to this problem. However the optimal decision rules could be characterized in a generic way as functions of the state variables using Bellman principle of optimality.

$$V(s_t) = \max_{\delta_t=(c_t,\ \alpha_t)\in D} u\left[s_t, \delta_t\right] + \beta \int V(s_{t+1}) p(ds_{t+1}|s_t, \delta_t) \tag{P1'}$$

with,

$$s_t = \left[W_t, \{nfi_s\}_{s=t,....H}, a_t, I_t)\right] \tag{12}$$

where W_t is financial wealth at period t,

$$W_t = A_{t-1} \times R(\alpha)_t \tag{13}$$

nfi_t is non financial income at period t

$$nfi_{it} = x'_{1it}\pi_1^{inc} + f_i\pi_2^{inc} + s\pi_3^{inc} + h_{it} \tag{14}$$

and a_t is the age of the agent

$$a_{t+1} = a_t + 1 \tag{15}$$

is an indicator function that equals 1 if the household had participated in the risky asset market in the previous period

$$I_t = \mathbf{1}\left(\alpha_{t-1} > 0\right) \tag{16}$$

3.3 Numerical solution to the agent problem

We need to solve the agent problem in order to evaluate her contribution to the pseudo-likelihood. The problem has not an analytical solution. Thus, we use numerical solution method to derive optimal decision rule for each individual.

We follow the method used by Imrohoroglu et al (1999) and also make use of the techniques proposed by Rust (1994 and 1996); Carroll (2002); Haliassos and Michaelides (2001) and Storesletten et al. (2007) in order to solve our dynamic structural model.

We proceed to discretize the space of the state variables into J-locations (called grid points) and find the optimal control variables and evaluate the value function for each point of the grid. Then, the optimal choice and the respective value of the value function for any value of the state variable is obtained interpolating results.

We solve the model using backward induction. At each period we use the Euler's conditions of the economic problem in order to obtain the optimal choice.

In the last period H the policy function is trivial (the agent consumes all her cash on hand) and the value function corresponds to the indirect utility function. Then, for each grid point we substitute this value function in the Bellman equation and compute the optimal policy rule for the previous period. The Euler's equations of the model are given by,

$$\mathrm{E}_t \left[u'(c_t) - \beta R(\alpha)_{t+1} u'(c_{t+1}) \right] = 0 \tag{17}$$

$$\mathrm{E}_t \left[A_t \ \left(R - R^f \right) u'(c_{t+1}) \right] = 0 \tag{18}$$

The model determines the optimal choice of consumption and risky asset share:

$$\{c_{it}^*, \alpha_{it}^*\} = \arg \max_{\{c, \alpha_{it}\} \in D} V(s_{it}) \tag{19}$$

$i = 1, 2, ...N$ denotes each individual in the sample and t the time period. It also yields $v_{it} V(s_{it} \mid \alpha_{it}^* \leqslant 0)$ the value function of period t under the assumption that the agent does not participate in the risky asset markets and $V(s_{it} \mid 0 < \alpha^*_{it} \leq 1)$ the value function of period t under the assumption that the agent participates in the risky asset markets. The solution method is detailed described in the Appendix A.

3.4 The policy rule

In the late sixties the seminal paper of Samuelson (1969) addressed the problem of consumption and portfolio choice over the life cycle by establishing a very simple rule in order to select the optimal portfolio. Assuming complete markets, independent and identically distributed returns, that household preferences can be represented by a power utility function and there are no other sources of income except financial returns, Samuelson showed that the sequence of portfolio structures that is statically optimal is also dynamically optimal. The optimal share of risky assets is constant, independent of wealth and age, and could be described as $\alpha^* = \frac{\mu}{\gamma\sigma^2}$, where α is the risky asset share of household´s portfolio, μ and σ^2 are, respectively, the expectation and variance of the excess return, and γ is the coefficient of relative risk aversion of the agent. Selection rule depends only on risk aversion, and the moments of the asset´s excess return distribution; that is to say that myopia is optimal and consequently life cycle patterns do not matter (Gollier, 2001).

Bodie, Merton and Samuelson (1992) introduce the presence of labor income into the picture and conclude that human capital could affect the decision about investment in the risky financial assets for two reasons: the uncertain and uninsurable nature of future labor income streams and the ability of households to vary their labor supply in the future.

This issue appear to be also very important from an empirical point of view because most of the observed volatility of households earnings comes from variations in labor incomes and typically risks related to human capital cannot be traded. Cocco *et al* (2005) show that the optimality of myopia disappears as the complete markets assumption is relaxed. These authors claim that the ratio of current wealth to expected future labor income is a crucial determinant of consumption and portfolio choice. To the extent that this ratio changes over the life cycle, the optimal portfolio allocation should not be expected to be age invariant and myopia could not be optimal. Other effects of labor income on portfolio choice appear if labor income and asset returns are correlated, or if we allow for the existence of time variation in the set of investment opportunities (Viceira, 2001).

Merton (1971) developed an optimal rule allowing for the existence of a constant labor income profile. Cocco *et al* (2005) describe Merton's rule and generalize it by allowing for variation in the future labor income realization. This rule could be described as $\alpha^* = \frac{\mu}{\gamma\sigma^2}\frac{W_t + PDV(FY_t)}{W_t}$ where W_t is financial wealth and $PDV(FY_t)$ is the expected present discounted value of the future labor income stream.

There are some professional financial planners who often advise that the fraction of wealth that people ought to hold in the risky asset markets should decline with age. A typical rule of thumb is that the percentage on an investor's portfolio of financial assets that is held in equities should equal 100 minus her age (Ameriks and Zeldes, 2000).

Viceira (2001) and Cocco *et al* (2005) solve numerically a life cycle model using calibration techniques, this method allows them to simulate the individual optimal path for consumption and risky asset holdings under the assumption that household receives an exogenous, uncertain and uninsurable stream of labor income. These authors conclude that the demand for the risky asset increases in the presence of labor income and that, conditional on a given future labor income stream, the optimal fraction to invest in the risky asset is a decreasing function of current wealth. They also find that the shape of the labor income profile over the life time ought to induce the investor to reduce the risky asset share when aging.

Figure 10 illustrates the shape of the optimal rule for the risky asset share depending on the financial wealth given the non-financial income stream by

Viceira, Cocco *et al* and this work. The difference between Viceira and Cocco *et al* is that the later work assumes that the agent is not allowed to borrow with the aim of investing in the risky asset markets. The difference between Cocco *et al* and our model is that we assume that the agent must pay certain non proportional costs to participate in the risky asset markets.

4 Estimation method

4.1 Indirect Inference

We estimate our structural dynamic programming model through indirect inference. Keane and Smith (2003) explain: "Indirect Inference exploits the ease and speed with which one can simulate data from complicated structural models. The basic idea of indirect inference is to view both the observed and the simulated data through the 'lens' of an descriptive statistical (or auxiliary) model characterized by a set of r parameters θ. The $k < r$ structural parameters β are then chosen so as to make the observed data and the simulated data look similar when viewed through this lens." In order to proceed with the Indirect Inference method we should specify both the auxiliary and the structural models.

4.2 The auxiliary statistical model

The actual data are on some of the state and all the control variables of the economic problem

$$\alpha_t = (\alpha_{1t}, \alpha_{2t}, ...\alpha_{Nt}) \quad X_t = (x_{1t}, x_{2t}, ...x_{Nt})$$

Here α_{it} and x_{it} are the risky asset share of household i and some observable characteristics of i in period t. The SHIW is a longitudinal dataset, thus we observe $\{\alpha_{it}, x_{it}\}_{i=1,...,N;\, t=1,....T.}$

The accurate specification of the auxiliary statistical model is crucial to the success of the econometric approach. The auxiliary statistical model specify a reduced-form model for the conditional probability distribution of the household shares of risky assets implicitly given by:

$$\alpha_{it} = \alpha^*_{it}\left[W_{it}, nfi_{it}, a_{it}, \alpha_{it-1}, Z_{it}\right] \tag{20}$$

The choice of the set of conditional variables is determined taking into account the state variables of the model, the expected sources of heterogeneity and the availability of information.

W_t is financial wealth, nfi_t is non-financial income, a_t is the age of the household's head, α_{it-1} is the household's risky asset share in the previous period and Z_t includes other characteristics of the household such as education, city size, working status and house ownership.

Most of the observed heterogeneity comes from the decision to participate in the risky asset markets. Thus, we focus on exploiting the information contained in the decision of participation. Therefore the dependent variable can assume only two values:

$$\delta_{it} = \begin{cases} 0 & \text{if } \alpha^*_{it} = 0 \\ 1 & \text{if } \alpha^*_{it} \in (0,1] \end{cases} \tag{21}$$

We use a logit model, thus the log-likelihood is given by:

$$\begin{aligned} L(\theta;\delta,X) &= \sum_{i=1}^{N} l_i(\theta) && (22)\\ l_i(\theta) &= \delta_{i1}\ln(p_i) + \delta_{i0}\ln(1-p_i) && (23)\\ p_i &= \Pr(\delta_i = 1 \mid x_i) = \Lambda(\mathbf{x}_i'\theta) && (24) \end{aligned}$$

where $\Lambda(\mathbf{r}) = \frac{\exp(\mathbf{r})}{1+\exp(\mathbf{r})}$.

4.3 The structural model

The structural model is given by:

$$\delta_i = \mathbf{1}\left[v_{i1} = \max(v_{i1}, v_{i0})\right]$$

where $v_{i0} = V(s_i \mid \delta_i = 0)$ the value function under the assumption that the agent does not participate in the risky asset markets and $v_{i1} = V(s_i \mid \delta_i = 1)$ the value function under the assumption that the agent participates in the risky asset markets. $s_i = (x_i, u_i)$ where x are observable variables, u unobservables, β a vector of structural parameters and δ the observed individual choice. The solution method was briefly described in Section 2 and it is thorough explained in Appendix A.

4.4 Generalized Indirect Inference

There are three different approaches in order to estimate a model using indirect inference. These three options differ in the choice of the metric that is used in order to measure the distance between the value of the estimations from the auxiliary model and those implied by the simulate techniques. Keane and Smith (2003) call these three ways Wald (W), Likelihood Ratio (LR) and Lagrange Multiplier (LM).

The approach we use here is the LR. The basic idea is to embed the solution of the structural model into the pseudo-likelihood of the auxiliary model.

The dependent variable of our model, δ_i, is a discrete random variable. Discrete random variables complicate the numerical calculation of the model because the objective surface is a step function. Step function arise when applying Indirect Inference to discrete choice models because any simulated choice δ_i^m is a step function of β (Altonji et al, 2006). To overcome this obstacle Keane and Smith (2003) propose to use a Generalized Indirect Inference approach, which consists in to smooth the objective surface. The key idea is to substitute the dependent variable of the model δ_i with a function of the latent utilities $\widetilde{\delta}_i$ as follows,

$$\widetilde{\delta}_i = \frac{1}{1 + \mathbf{exp}\left[\frac{v_{i0} - v_{i1}}{\lambda}\right]} \tag{25}$$

Note that if $v_{i0} = \max(v_{i1}, v_{i0})$ and $\lambda \to 0$ then $\widetilde{\delta}_i \to \delta_i = 0$ and if $v_{i1} =$

$\max(v_{i1}, v_{i0})$ and $\lambda \to 0$ then $\widetilde{\delta}_i \to \delta_i = 1$.

For each individual we observe the cash on hand at period t (coh_{it}) , non-financial income (nfi_{it}) for periods $t-s, t-(s-1), ...t.$, the agent's age (a_{it}) and the participation in the risky asset market in the previous period $(\mathbf{I}_{it-1})$

Given that information, a prediction model and realizations from a multivariate $\mathcal{N}(0, I)$ we simulate a random process for the future income stream $\{nfi_{is}^m\}_{i=1...N,\ s=t+1,...T}$. Then we solve the economic problem of the agent for a given value of the structural coefficients (β) compute $v_{ij} = V(s_i, \beta \mid \delta_i)$ $j = 0, 1;$ and obtain

$$\widetilde{\delta}^m(\beta) = \left(\widetilde{\delta}_1^m(\beta), ..., \widetilde{\delta}_N^m(\beta)\right)$$

Afterwards we estimate

$$\widetilde{\theta}^m(\beta) = \arg\max_{\theta} L(\theta; \widetilde{\delta}^m(\beta), X) \tag{26}$$

$$L(\theta; \widetilde{\delta}^m(\beta), X) = \sum_{i=1}^{N} \widetilde{\delta}_i^m(\beta) \log \Lambda(x_i'\theta) + \left(1 - \widetilde{\delta}_i^m(\beta)\right) \log\left[1 - \Lambda(x_i'\theta)\right]$$

where $\Lambda(\mathbf{r}) = \frac{\exp(\mathbf{r})}{1+\exp(\mathbf{r})}$.

We repeat that procedure for $m = 1, 2, ...M$ and compute

$$\widetilde{\theta}(\beta) = \frac{1}{M} \sum_{m=1}^{M} \widetilde{\theta}^m(\beta) \tag{27}$$

Finally the estimation for the structural parameters β comes from

$$\widehat{\beta} = \arg\max_{\beta} L(\widetilde{\theta}(\beta); \delta, X)$$

where

$$L(\widetilde{\theta}(\beta); \delta, X) = \sum_{i=1}^{N} \delta_i \log \Lambda\left(x_i'\widetilde{\theta}(\beta)\right) + (1 - \delta_i) \log\left[1 - \Lambda\left(x_i'\widetilde{\theta}(\beta)\right)\right] \tag{28}$$

where $\Lambda(\mathbf{r}) = \frac{\exp(\mathbf{r})}{1+\exp(\mathbf{r})}$.

4.5 Standard errors

The log-likelihood of the auxiliary model $L(\theta;\delta,X)=\sum_{i=1}^{N} l_i\left(\theta\right)$ where θ is $r\times 1$. Let $\widehat{\theta}=\arg\max_{\theta} L(\theta;\delta,X)$ and such that plim $\widehat{\theta}=\theta_0$. The robust estimation of its asymtotic variance is given by

$$\widehat{Var}(\widehat{\theta})=\frac{1}{N}\widehat{H}^{-1}\widehat{W}\widehat{H}^{-1} \tag{29}$$

where $\widehat{H}$ and $\widehat{W}$ are consistent estimations of

$$H = \text{p}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{\partial^2 l_i\left(\theta_0\right)}{\partial\theta\partial\theta'} \tag{30}$$

$$W = \text{p}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{\partial l_i\left(\theta_0\right)}{\partial\theta}\frac{\partial l_i\left(\theta_0\right)}{\partial\theta'} \tag{31}$$

To the extend that $L(\theta;\delta,X)$ is a pseudo-likelihood, the identity of the information matrix does not hold, in general $W\neq -H$.

Notice that the auxiliary model is a logit model, thus we have:

$$\widehat{H} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\lambda}_i x_i x_i^{'} \tag{32}$$

$$\widehat{W} = \frac{1}{N}\sum_{i=1}^{N}\left(\delta_i-\widehat{\Lambda}_i\right)^2 x_i x_i^{'} \tag{33}$$

where $\widehat{\Lambda}_i=\frac{\exp\left(x_i^{'}\widehat{\theta}\right)}{1+\exp\left(x_i^{'}\widehat{\theta}\right)}$ and $\widehat{\lambda}_i=\widehat{\Lambda}_i\left(1-\widehat{\Lambda}_i\right)$.

Let β a $k\times 1$ vector of the structural model's parameters. The variance of $\widetilde{\beta}=\arg\max_{\beta} L(\widetilde{\theta}(\beta);\delta,X)$ is given by

$$\widehat{Var}(\widehat{\beta})=\frac{1}{N}\left(\widehat{D}'\widehat{H}\widehat{D}\right)^{-1}\widehat{D}'\widehat{W}\widehat{D}\left(\widehat{D}'\widehat{H}\widehat{D}\right)^{-1} \tag{34}$$

where $\widehat{D}$ is the $r\times k$ matrix of the numerical partial derivatives evaluated at $\widehat{\beta}$:

$$\widehat{D}=\frac{\partial\widetilde{\theta}(\widehat{\beta})}{\partial\beta'} \tag{35}$$

5 Returns, preference parameters and non-financial income predictions

5.1 Asset returns and preference parameters

The investment opportunity set is constant and there are two financial assets, one risk-less asset (cash and deposits) and another one risky (stocks and bonds).

In order to specify the return of the risky asset we assume that there are only two states of nature, bad and good, each with the same probability (50 percent). The risky asset gross return (denoted by R_s) is given by $\underline{R}$ and $\overline{R}$ in the bad and good states respectively.

$$R_s = \begin{cases} \underline{R} & \Pr = 0.5 \\ \overline{R} & \Pr = 0.5 \end{cases} \tag{36}$$

We use the data in Pelizzon and Weber (2007) and compute both the expected value and the deviation of the excess return for the efficiency portfolio. $\underline{R}$ and $\overline{R}$ are fixed at 0.86 and 1.25 respectively. The source data are reported in Table 3.

Risk-less asset yields a constant gross after tax real return, R^f. R^f is fixed in 1.02 following the estimation of the bank deposit rates in Italy by Panetta and Violi (1999) for the period 1981-1994.

Preference parameters are fixed at plausible values. We chose the subjective discounted value factor $B = 0.96$ and the coefficient of relative risk aversion $crra = 3$.

5.2 Prediction of non-financial income process

The nonfinancial income of the household is assumed to be governed by the following stochastic process,

$$nfi_{is} = x'_{1is}\pi_1^{inc} + f_i\pi_2^{inc} + t\pi_3^{inc} + h_{is} \qquad s = t, t+1, ...H \tag{37}$$

where

nfi_{is} : nonfinancial income of the household

x'_{1is} : age and age square of the household' head

η_i : an unobservable time invariant individual characteristic

f'_i :dummies of household's head education level, a dummy indicating if the household resides in the South of Italy and dummies indicating the city size.

t a polynomial in time and a dummy to 1998.

h_{is} is a stochastic term

The first step is to estimate the following OLS equation:

$$\begin{aligned} nfi_{it} &= x'_{1it}\pi_1^{inc} + f_i\pi_2^{inc} + t\pi_3^{inc} + h_{it} \\ &= w'_{it}\pi^{inc} + h_{it} \end{aligned} \tag{38}$$

It is important to remark that the unique goal of this model is to simulate the future nonfinancial income stream for each household in the sample, from her current age to the maximum age allow in the model (H). Consequently the variables to be included in the model either should follow a deterministic low of motion or must being able to be treated as if they were time invariant. For this reason some variables, which are known to be significant for the income process, are not taking it into account (e.g. industry, occupation, family size, percent of household members in the labor force, etc.). Estimations from the OLS regression are reported in Table 4.

Notice also that h_{it} is assumed to be uncorrelated with regressors not because that is true but cause the focus here is not about π^{inc} coefficients itself. The model for the prediction of the nonfinancial income stream is completely specificated by adding the following assumptions

$$\begin{aligned} h_{it} &= \eta_{1i} + v_{it} & (39) \\ v_{it} &= \sum_{s=1}^{q} \rho_s^v v_{it-s} + \varepsilon_{it} \qquad \|\rho^v\| < 1 & (40) \\ \varepsilon_{it}^2 &= \sum_{s=1}^{p} \rho_s^\varepsilon \varepsilon_{it-s}^2 + \eta_{2i} + \xi_{it} \qquad \|\rho^\varepsilon\| < 1 & (41) \\ \xi_{it} &\sim NIID(0, \sigma_\xi^2) \end{aligned}$$

The persistence coefficients ρ_s^v are consistent estimated by using the Arellano-Bond GMM estimator (Arellano and Bond, 1991; Arellano, 2003). Results are presented in Table 5. From it we conclude that an AR(2) model is the most accurate in order to capture the non-financial income dynamics.

Afterwards, the estimation of the individual fixed effect of the non-financial income process η_{1i} is simply obtained as:

$$\begin{aligned} \widehat{\eta}_{1i} &= \frac{1}{T-q} \sum_{t=q}^{T-q} \left(nfi_{it} - w_{it}' \widehat{\pi}^{inc} - \sum_{s=1}^{q} \widehat{\rho}_s^v \widehat{\nu}_{it-s} \right) & (42) \\ \widehat{\nu}_{it-s} &= nfi_{it-s} - w_{it-s}' \widehat{\pi}^{inc} & (43) \end{aligned}$$

Table 6 shows some measure of the estimated model. Some conclusions can be inferred from it. First, the most important components of non-financial income are the time invariant characteristics of the agent. Second, both the time invariant observable characteristics and the individual fixed effect account for most of the volatility of the non-financial income. Third, estimated idiosyncratic shocks and the unobservable fixed effect are highly correlated.

Finally, an Arch(p) model for the variance of the non-financial income process is estimated. The coefficients ρ_s^ε are obtained using Arellano-Bond GMM estimator. We also estimate an individual fixed effect in the variance of the non-financial income model. We calculate it as follows,

$$\widehat{\eta}_{2i} = \frac{1}{T-p-1} \sum_{t=p+1}^{T-p-1} \left(\widehat{\varepsilon}_{it}^2 - \sum_{s=1}^{p} \widehat{\rho}_s^{\varepsilon} \widehat{\varepsilon}_{it-s}^2 \right) \tag{44}$$

From Table 7 we can draw the conclusion that there is not an Arch process associated with the analyzed non-financial income process. However, as well as the $\widehat{\eta}_{1i}$ in the prediction of the non-financial income level the estimated individual fixed effect $\widehat{\eta}_{2i}$ for the variance of the process is very important in order to fit the non-financial income stream as close as possible. Then, $\widehat{\eta}_{2i}$ is simple calculated as $\frac{1}{T} \sum_{t=1}^{T} \left(\widehat{\varepsilon}_{it}^2 \right)$.

Summarizing, the non-financial-income stream predictions are calculated using the following three equations,

$$\widehat{nfi}_{is} = x'_{1is} \widehat{\beta}_1^{inc} + f'_{i2} \widehat{\beta}_2^{inc} + t \widehat{\delta}^{inc} + \widehat{\eta}_{1i} + \widehat{v}_{is} \tag{45}$$

$$\widehat{v}_{is} = \sum_{q=1}^{2} \widehat{\rho}_s \widehat{v}_{is-q} + \widetilde{\varepsilon}_{is} \tag{46}$$

$$\widetilde{\varepsilon}_{is} = \widetilde{z}_{is} \times \sqrt{\widehat{\eta}_{2i}} \tag{47}$$

where $s = t+1, \ldots.., H$ and $\widetilde{z}_{is}$ is a pseudo-random number generated from the normal standard distribution.

6 Results

6.1 Simulations for the policy function

Some simulations are made in order to illustrate the yields of the model. Returns and preference parameters as well as the simulation of the non-financial income of each individual in the sample are as it was described in Section 5. Participation cost is fixed at the bottom bound of 0.175 percent of the non-financial income. Figures 11 to 16 show some measures of our simulated results. We focus on the relationship between the simulated risky asset shares and the main state variables of the economic problem (age, cash-on-hand/non-financial income ratio and non-financial income).

Figure 11 illustrates the relationship between the simulated risky asset shares and the age of the agent. It shows that the model is able to generate heterogeneous results. Moreover, it demonstrates that, in the presence of non proportional participation costs zero solution could be an optimal response. Figure 12 plots the average of the optimal risky asset shares against the age and suggests that the age of the agent does not matter in order to select the optimal portfolio. This result differs than that of Cocco *et al* (2005), Viceira (2001) and Jagannathan and Kocherlakota (1999). These authors have found some support for the professional financial planners who often advise that the fraction of wealth that people ought to hold in the risky asset markets should decline with age.

Figure 13 deals with the relationship between the simulated risky asset shares and the cash-on-hand/non-financial income ratio. A highly non-linear relationship is found there. Results seem to mimic the expected optimal policy rule showed in Figure 10. Thus, our results are in line with those of the recent literature on the household portfolio choice (see, e.g., Cocco *et al, 2005;* Viceira, 2001; Campbell, 2006). However, Figure 13 shows that the zero solution could be optimal even for large values of the cash-on-hand/non-financial income ratio. On the other hand, when analyzing Figure 14 it is easy to conclude that the optimal risky asset share is an increasing function of the cash-on-hand/non-financial income ratio. In particular, it appears that to the majority of those on the bottom of the cash-on-hand/non-financial income ratio distribution is optimal not to invest in the risky asset markets even if the participation cost is very small.

We analyze how does the optimal risky asset share varies with non-financial income of the agent in Figures 15 and 16. It is important to notice that in these figures the independent variable is the observed non-financial income of the agent in the period t and not the present discounted value of the non-financial income stream. Once again in Figure 15 we observe a lot of heterogeneity in the optimal risky asset shares and find that zero solution could be an optimal choice all over the non-financial income distribution. Finally, as it was expected, Figure 16 shows that the average of the optimal risky asset shares is an increasing function of non-financial income.

6.2 Empirical Results

We estimate a simpler version of the model assuming that the non proportional participation costs are given exclusively by a single opportunity cost. This cost depends on the individual level of non-financial income of the period, that is, it could be modelled as $G_{it} = g \times nfi_{it}$. That is to say, our structural model includes only one structural parameter: g. We allow g to varies among groups. In particular, we split the sample into four categories depending upon the education level of the household's head. These four categories are: elementary, secondary, high school and college or post-graduated.

We proceed first with the estimation of the auxiliary model using actual data: we obtain $\widehat{\theta}$.[2] Then we estimate M times the auxiliary model using our M simulated samples: we collect $\widetilde{\theta}^m(\beta)$ and therefore we can calculate $\widetilde{\theta}(\beta)$. Afterwards we evaluate the objective function of the problem (given by equation [28]) and found the value of the parameter β that maximize its value. Finally, we calculate the standard error of the estimation by evaluating the equation [34].

We found that the participation costs are always positive and significant but strong differences among education groups are observed. As expected participation cost is lower the higher is the education level of the agent. Table 9 reports the estimation of the participation costs for each education group.

The first line of Table 9 includes the results for the individuals graduated

[2] In the Table 8 we report the estimation of the auxiliary model for each education group.

or post-graduated. The sample size is 270. The estimated cost for this group is very low: attains to 0.175 percent of the non-financial income. However, it is positive and significant. Table 10 shows that in terms of money this cost represents 103 euros of 2006.

In the second line we find the result for the high school sample. There are 964 individuals in this group. The estimation of the cost to participate is also significant and it is substantially greater than that we estimate for college and postgraduated. In term of non-financial income the estimated cost is 1.65 percent, equivalent to 719 euros of 2006.

The group of secondary school includes 926 individuals. The estimation of the participation cost for this group in term of the non-financial income doubles the respective for the high school group: attains to 3.35 percent. However, the difference in the equivalence in euros is smaller: 1157 against 719 euros of 2006.

Finally, with a sample of 795 individuals the cost of participating in the risky asset markets for the elementary school group is estimated as high as the 6 percent of the non-financial income. Curiously the equivalence in euros attains to 1126, very similar than that obtained for the group of elementary school.

These results are comparable with those Paiella (2006) and Attanasio and Paiella (2006) found using US data. These authors estimated respectively an average lower bound ranging 0.7 and 3.3 percent and a 0.4 percent of non-durable consumption.[3]Paiella (2006) has also established a lower bound of the participation cost at 130 dollars.

7 Concluding remarks

This paper develops and estimates a dynamic structural model of participation in the risky asset market using household level panel data. The data

[3]In the SHIW survey the non-durable consumption represents the 70 percent of non-financial income.

come from the Italian Survey of Households Income and Wealth.

The model we use is the simplest one we can imagine in order to capture the portfolio choice over the life cycle. We proceed by assuming that households choose the optimal portfolio and calculate how much cost is needed in order to explain the observed households' behaviour. We solve the model using numerical techniques. Then we embed the optimal solution into the statistical (auxiliary) model in order to estimate the structural parameters using Generalized Indirect Inference.

Our simulated results support some conclusions of the recent literature on the issue of portfolio choice over the life cycle. The optimal risky asset share is an increasing function of both the cash-on-hand/non-financial income ratio and the non-financial income. On the contrary, our results do not exhibit any systematic relation with respect to the age of the agent.

This paper focuses on the estimation of the non proportional costs to participate in the risky asset market. All the remaining parameters of the model (asset returns, preference parameters, non-financial income process) are either fixed or estimated outside the core estimation model.

We consider heterogeneous costs among education groups. The education group is determined by the education level of the household's head. We distinguish four groups: elementary school, secondary school, high school and college or post-graduated.

We found that participation costs in the risky asset markets are positive and significant. We also conclude that they vary a lot with the education level of the household head: the estimated cost for college and post-graduated is 0.175 percent of non-financial income but rises to a 6 percent for the elementary school group. The results for secondary and high school groups attain to 3.35 and 1.65 percent of non-financial income respectively. In term of euros of 2006 the estimated costs attain to 103, 719, 1157 and 1126 for the college and post-graduated, the high school, the secondary and the elementary groups respectively. These results are comparable with those Paiella (2006) and Attanasio and Paiella (2006) found using US data.

The model could be extended in order to analyze and estimate other aspects of the problem. For example, it could be easily changed in order to estimate how does the participation cost vary among other sources of heterogeneity like geographic regions, previous participation in the risky asset markets. It could be also interesting to estimate the evolution of these costs over time. Furthermore, we could analyze how sensitive are the results to variation on parameters like the coefficient of risk aversion or the asset returns. Some more difficult extensions could also be done. In particular, the introduction of different types of costs such opportunity and monetary, entry and monitoring. Finally, some ambitious routes could be thought like introducing housing or changing the assumptions about the utility function.

References

[1] Ameriks, J and S.P. Zeldes (2000), "How Do Household Portfolio Shares Vary with Age", mimeo.

[2] Arellano, M. (2003), *Panel Data Econometrics,* Oxford University Press: Advanced Texts in Econometrics.

[3] Arellano, M. and S. Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 58, 277-297.

[4] Attanasio, O. and M. Paiella (2006), "Intertemporal Consumption Choices, Transaction Costs and Limited Participation to Financial Markets: Reconciling Data and Theory", NBER WP 12412.

[5] Bodie, Z., Merton, R.C. and W.F. Samuelson (1992), "Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model", *Journal of Economic Dynamics and Control*, 16, 427-449.

[6] Campbell, J. (2006), "Household Finance", *Journal of Finance,* 61, pp. 1553-1604.

[7] Campbell, J. and L. Viceira (2002), *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors,* Oxford University Press, New York.

[8] Carroll, C. (2002), *Lecture Notes on Solution Methods for Microeconomic Dynamic Stochastic Optimization Problems,* Lecture Notes, John Hopkins University.

[9] Cocco, J. (2004), "Portfolio Choice in the Presence of Housing", *Review of Financial Studies,* 18, 535-567.

[10] Cocco, J.; Gomes, F. and P. Maenhout (2005), "Consumption and Portfolio Choice over the Life-Cycle", *Review of Financial Studies,* 18, 491-533.

[11] Deaton, Angus (1991), "Saving and Liquidity Constraints", *Econometrica,* 59, 1221-1248.

[12] Gollier, Christian (1999), "What does the Classical Theory Have to Say about Household Portfolios?", in: L. Guiso, M. Haliassos and T. Japelli, eds., *Household Portfolios* (MIT Press)..

[13] Guiso, L. and T. Jappelli (2001), "Household Portfolios in Italy", in: L. Guiso, M. Haliassos and T. Japelli, eds., *Household Portfolios* (MIT Press).

[14] Haliassos M. and Michaelides A. (2001), "Calibration and Computation of Household Portfolio", in: L. Guiso, M. Haliassos and T. Japelli, eds., *Household Portfolios* (MIT Press).

[15] Heaton, J. and D.J. Lucas (2000), "Portfolio choice and asset prices: The Importance of Entrepreneurial Risk", *Journal of Finance, 55, 1163-1198.*

[16] Imrohoroglu, A.; Imrohoroglu, S. and D.H. Joines (1999), "Computing Models of Social Security", in: R. Marimon and A. Scott, eds., *Computational Methods for the Study of Dynamics Economies* (Oxford University Press).

[17] Jagannathan, R. and N. Kocherlakota (1996), "Why should Older People Invest Less in Stocks than Younger People", Federal Reserve Bank of Minneapolis *Quarterly Review,* 20, 11-23.

[18] Merton, R. C. (1971), "Optimum Consumption and Portfolio Rules in a Continous-Time Model", *Journal of Economic Theory*, 3, 373-413.

[19] Miniaci, R. and G. Weber (2001), "Econometric Issues in the Estimation of Household Portfolio Models", in: L. Guiso, M. Haliassos and T. Japelli, eds., *Household Portfolios* (MIT Press).

[20] Paiella, M. (2006), "The Foregone Gains of Incomplete Portfolios", CSEF Working Paper No. 156.

[21] Panetta, F. and R. Violi (1999), "Is there an Equity Premium Puzzle in Italy? A Look at Asset Returns, Consumption and Financial Structure Data Over the Last Century", Banca d'Italia, Temi di discussione No. 353.

[22] Pelizzon, L. and G. Weber (2007), "Efficient Portfolios when Housing Needs Change over the Life-Cycle", mimeo.

[23] Perraudin, W.R.M. and B.E. Sorensen (2000), "The Demand for Risky Assets: Sample Selection and Household Portfolios", *Journal of Econometrics,* 97, 117-144.

[24] Rust, John (1996), "Numerical Dynamic Programming in Economics", in: H.M. Amman, D.A. Kendrick and J. Rust, eds., *Handbook of Computational Economics*, Vol. 1 (Elsevier Science, Amsterdam).

[25] Rust, John (1994), "Structural Estimation of Markov Decision Processes", Handbook of Econometrics, Vol. 4, Chapter 51.

[26] Samuelson, Paul A. (1969), "Lifetime Portfolio Selection by Dynamic Stochastic Programming", *Review of Economics and Statistics*, 51, 239-246.

[27] Storesletten, K.; Telmer, C.L. and A. Yaron (2007), "Asset Pricing with Idiosyncratic Risk and Overlapping Generations", *Review of Economic Dynamics,* forthcoming.

[28] Viceira, Luis M. (2001), "Optimal Portfolio Choice for Long-Horizon Investor with Nontradable Labor Income", *Journal of Finance,* LVI(2), 433-470.

[29] Vissing-Jorgensen (2002), "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution", Journal of Political Economy, 110, 825-853.

Appendix A. Numerical Solution for the agent problem

The problem can not be solved analytically. We use numerical solution method to derive optimal decision rule for each individual. We proceed to discretize the space of the state variables and find the optimal control variables for each value of them within a grid for the state variables. Then, for each specific value of the grid for the state variables we obtain the optimal decision rule and then compute the optimal choice for any value of the state variable interpolating results.

We solve the model using backward induction. At each period we use the Euler condition of the economic problem in order to obtain the optimal choice.

In the last period the policy functions are trivial (the agent consumes all her cash on hand) and the value function corresponds to the indirect utility function. Then, for each grid point we substitute this value function in the Bellman equation and compute the optimal policy rule for the previous period. We optimize using the Euler equation of the model given by,

$$\mathrm{E}_t\left[u'\left(c_t\right) - \beta R\left(\alpha\right)_{t+1} u'\left(c_{t+1}\right)\right] = 0 \tag{48}$$

$$\mathrm{E}_t\left[A_t\ \left(R - R^f\right) u'\left(c_{t+1}\right)\right] = 0 \tag{49}$$

at s=H

The solution is trivial, the agent will consume all her cash on hand:

$$c_H = coh_H = W_H + nfi_H \tag{50}$$

$$V\left(s_H\right) = u(c_H) \tag{51}$$

The numerical solution is obtained over a finite number of J-locations (called grid points) for coh_H

at s=H-1

We build a J-grid points for cash–on-hand at period $H-1$ and for each value in the grid the problem is solved by minimizing

$$\begin{aligned} sqe_{H-1} &= e^2_{1H-1} + e^2_{2H-1} && (52) \\ e_{1H-1} &= E_{H-1}\left[u'\left(c_{H-1}\right) - BR\left(\alpha_{H-1}\right)u'\left(coh_H\right)\right] && (53) \\ e_{2H-1} &= E_{H-1}\left[A_{H-1}\ \left(R - R^f\right)u'\left(coh_H\right)\right] && (54) \\ coh_H &= W_H + nfi_H && (55) \end{aligned}$$

where

$$W_H = \begin{cases} A_{H-1} \times R^f & \text{if } \alpha_{H-1} = 0 \\ A_{H-1} \times R_j(\alpha_{H-1}) \quad \text{j=low, high} & \text{if } \alpha_{H-1} > 0 \end{cases} \qquad (56)$$

with

$$\begin{aligned} A_{H-1} &= \left(coh_{H-1} - G_{H-1} - c_{H-1}\right) && (57) \\ coh_{H-1} &= W_{H-1} + nfi_{H-1} && (58) \\ G_{H-1} &= \begin{cases} 0 & if\ \alpha_{H-1} = 0 \\ g_E \times nfi_{H-1} & if\ \alpha_{H-1} > 0 \text{ and } I_{H-1} = 0 \\ g_M \times nfi_{H-1} & if\ \alpha_{H-1} > 0 \text{ and } I_{H-1} = 1 \end{cases} && (59) \end{aligned}$$

In order to obtain $V(W_H, nfi_H, H, I_H)$ we use piecewise linear interpolation (outside the grid we do linear extrapolation). Then we obtain $V(W_{H-1}, nfi_{H-1}, H-1, 0)$ and $V(W_{H-1}, nfi_{H-1}, H-1, 1)$ for each grid point of $coh_j = W^j_{H-1} + nfi_{H-1}$

at s=t+1,...H-2

At each period s from t+1 to H-2 the solution for each grid point $coh^j_s = W^j_s + nfi_s$ is attained by finding $\delta_s = (c_s, \alpha_s)$ that minimize the equation

$$sqe_s = e^2_{1s} + e^2_{2s} \qquad (60)$$

$$\begin{aligned} e_{1s} &= E_s\left[u'\left(c_s\right) - BR\left(\alpha\right)u'\left(c_{s+1}\right)\right] && (61) \\ e_{2s} &= E_s\left[A_s\ \left(R - R^f\right)u'\left(c_{s+1}\right)\right] && (62) \end{aligned}$$

with

$$
\begin{aligned}
c_{s+1} &= \arg\max_{c_{s+1}} V(s_{s+1}) && (63)\\
W_{s+1} &= \begin{cases} A_s \times R^f & \text{if } \alpha_s = 0 \\ A_s \times R_j(\alpha_s) \quad \text{j=low, high} & \text{if } \alpha_s > 0 \end{cases} && (64)\\
A_s &= coh_s - G_s - c_s && (65)\\
G_s &= \begin{cases} 0 & if\ \alpha_s = 0 \\ g_E \times nfi_s & if\ \alpha_s > 0 \text{ and } I_s = 0 \\ g_M \times nfi_s & if\ \alpha_s > 0 \text{ and } I_s = 1 \end{cases} && (66)
\end{aligned}
$$

both optimal decision rule $\delta_s = (c_s, \alpha_s)$ and the respective value function $V(W_s, nfi_H, H, I_H)$ are obtained by piecewise linear interpolation (outside the grid we do linear extrapolation).

at s=t

At period t the solution for a given value of $W_t, \{nfi_s\}_{s=t,....H}, a_t, I_t$, is attained by finding $\delta_t = (c_t, \alpha_t)$ that minimize the equation

$$sqe_t = e_{1t}^2 + e_{2t}^2 \tag{67}$$

$$
\begin{aligned}
e_{1t} &= E_t\left[u'(c_t) - BR(\alpha_t)\, u'(c_{t+1})\right] && (68)\\
e_{2t} &= E_t\left[A_t\ \left(R - R^f\right) u'(c_{t+1})\right] && (69)
\end{aligned}
$$

with

$$
\begin{aligned}
\{c_{t+1}, \alpha_{t+1}\} &= \arg\max_{c_{t+1},\alpha_{t+1}} V(s_{t+1}) && (70)\\
W_{t+1} &= \begin{cases} A_t \times R^f & \text{if } \alpha_t = 0 \\ A_t \times R_j(\alpha_t) \quad \text{j=low, high} & \text{if } \alpha_t > 0 \end{cases} && (71)\\
A_t &= coh_t - G_t - c_t && (72)\\
G_t &= \begin{cases} 0 & if\ \alpha_t = 0 \\ g_E \times Y_t & if\ \alpha_t > 0 \text{ and } I_t = 0 \\ g_M \times Y_t & if\ \alpha_t > 0 \text{ and } I_t = 1 \end{cases} && (73)
\end{aligned}
$$

The model yields $V(s_t \mid \alpha^*_{it} \leqslant 0)$ and $V(s_t \mid 0 < \alpha^*_{it} \leq 1)$ which determines:

$$\alpha_{it} = \mathbf{1}\left[v_{it1} = V(s_t \mid \alpha^*_{it} > 0) = \max\left(v_{it0}, v_{it1}\right)\right] \tag{74}$$

$i = 1, 2, ...N$ denotes each individual in the sample and t the time period.

Table 1. Descriptive statistics for the whole sample

	Total	***1987***	***1989***	***1991***	***1993***	***1995***	***1998***	***2000***	***2002***	***2004***
Age										
Average	52.6	52.0	49.8	51.2	52.2	52.5	52.6	54.3	55.6	56.5
Standard deviation	14.0	14.8	13.4	13.5	13.9	13.7	13.4	13.8	13.6	13.9
Non-financial income (euros, prices 1995)										
Average	25,230	23,588	24,742	24,316	24,513	24,623	25,910	26,648	27,458	27,995
Standard deviation	17,453	17,669	14,786	13,885	15,683	17,839	18,588	17,967	20,390	21,061
Financial Assets (euros, prices 1995)										
Average	19,841	16,521	15,729	14,845	19,875	20,768	25,163	26,692	24,297	21,187
Standard deviation	60,158	33,008	43,516	35,958	56,178	54,124	82,305	99,487	74,339	63,160
Real Wealth (euros, prices 1995)										
Average	126,781	83,762	80,643	108,371	139,807	141,072	146,947	151,233	167,154	190,417
Standard deviation	261,893	263,010	133,651	157,430	252,075	237,002	382,217	279,621	296,580	315,849
Risky assets										
Proportion of participants	0.32	0.24	0.25	0.31	0.37	0.40	0.34	0.36	0.34	0.33
Education										
Elementary school	0.30	0.33	0.33	0.34	0.31	0.30	0.25	0.27	0.28	0.27
Middle School	0.29	0.25	0.29	0.30	0.31	0.29	0.29	0.29	0.30	0.30
High school	0.26	0.22	0.24	0.23	0.24	0.27	0.33	0.31	0.30	0.31
Bachelor's degree + post-graduated	0.08	0.11	0.08	0.08	0.07	0.07	0.09	0.09	0.08	0.08
City size										
From 5 to 20 thousands	0.14	0.13	0.13	0.12	0.13	0.17	0.16	0.16	0.15	0.17
From 20 to 50 thousands	0.29	0.28	0.25	0.29	0.28	0.32	0.33	0.31	0.30	0.27
From 50 to 200 thousands	0.30	0.33	0.26	0.29	0.35	0.28	0.28	0.31	0.30	0.30
More than 200 thousands	0.17	0.19	0.28	0.22	0.15	0.13	0.13	0.12	0.11	0.12
Resident in the South	0.36	0.35	0.39	0.40	0.37	0.36	0.37	0.34	0.33	0.32
Working status										
Employed	0.64	0.69	0.73	0.68	0.63	0.61	0.62	0.58	0.54	0.52
Houseown										
Proportion of owners	0.49	na	na	0.66	0.68	0.71	0.73	0.74	0.75	0.75
Number of observations	43730	7649	5725	5396	4781	4816	4167	3935	3700	3561

Table 2. Descriptive statistics for the used sample

	Total	***1993***	***1995***	***1998***	***2000***	***2002***	***2004***
Age							
Average	57.1	53.9	54.8	55.6	57.1	58.2	59.5
Standard deviation	12.3	12.4	12.3	12.0	11.9	12.1	12.5
Non-financial income (euros, prices 1995)							
Average	27,611	24,123	25,666	27,192	28,869	28,442	27,784
Standard deviation	14,095	13,826	13,885	13,820	14,628	14,065	13,906
Financial Assets (euros, prices 1995)							
Average	26,890	14,472	25,597	27,424	33,661	26,821	22,336
Standard deviation	61,590	16,639	52,616	57,001	94,819	43,479	43,032
Real Wealth (euros, prices 1995)							
Average	168,939	116,789	160,240	149,465	157,983	191,903	195,403
Standard deviation	238,485	136,738	208,977	228,340	167,306	315,554	262,338
Risky assets							
Proportion of participants	0.44	0.36	0.47	0.45	0.48	0.44	0.40
Education							
Elementary school	0.26	0.39	0.27	0.26	0.24	0.24	0.27
Middle School	0.30	0.21	0.30	0.30	0.31	0.32	0.30
High school	0.31	0.24	0.31	0.31	0.33	0.33	0.31
Bachelor's degree + post-graduated	0.09	0.11	0.08	0.08	0.09	0.08	0.09
City size							
From 5 to 20 thousands	0.16	0.09	0.14	0.14	0.16	0.15	0.20
From 20 to 50 thousands	0.30	0.31	0.26	0.32	0.29	0.29	0.30
From 50 to 200 thousands	0.33	0.40	0.40	0.32	0.32	0.33	0.31
More than 200 thousands	0.12	0.13	0.15	0.13	0.13	0.12	0.08
Resident in the South	0.34	0.55	0.42	0.37	0.32	0.31	0.29
Working status							
Employed	0.52	0.63	0.58	0.56	0.52	0.50	0.45
Houseown							
Proportion of owners	0.79	0.75	0.78	0.77	0.79	0.80	0.82
Number of observations	3061	87	368	770	645	491	700

Table 3: First and second moments of annual excess returns in Italy (1989-2003)

	Government Bonds	Corporate Bonds	Stocks
Expected return %	4.0981	2.2845	4.9011
Standard Deviation %	5.2383	3.2169	28.995

Correlation	Government Bonds	Corporate Bonds	Stocks
Government Bonds	1	0.8404	0.0215
Corporate Bonds		1	0.1726
Stocks			1
Weights of the efficiency portfolio	0.67	0.29	0.04

Source: Pelizzon and Weber (2007)

Table 4: OLS Regression
Dependent variable: Nonfinancial-income (log)

	Model 1		*Model 2*	
	Coef.	***Std. Err.***	***Coef.***	***Std. Err.***
Age	0.032	(0.001)**	0.064	(0.001)**
Age square	-0.001	(0.000)**	-0.002	(0.000)**
Time	0.309	(0.021)**	0.351	(0.027)**
Time square	-0.321	(0.025)**	-0.338	(0.031)**
Time third power	0.094	(0.008)**	0.094	(0.010)**
Dummy 1998	-0.006	(0.007)	0.004	(0.008)
Elementary school	0.179	(0.008)**	0.181	(0.011)**
Middle school	0.346	(0.009)**	0.367	(0.011)**
High school	0.493	(0.009)**	0.598	(0.011)**
More than high school	0.723	(0.012)**	0.929	(0.013)**
Resident in the South	-0.231	(0.004)**	-0.298	(0.005)**
City size (5 to 20 th.)	-0.019	(0.007)**	-0.025	(0.009)**
City size (20 to 50 th.)	-0.001	(0.007)	-0.005	(0.008)
City size (50 to 200 th.)	0.009	(0.007)	-0.003	(0.008)
City size (More than 200 th.)	0.030	(0.007)**	0.012	(0.009)
Family size	0.039	(0.002)**		
#Labor income earners	0.324	(0.003)**		
Constant	7,924	(0.018)**	8,552	(0.015)**
Occupation dummies				
F(7,43639)	154.8			
Industry dummies				
F(8,43639)	237.2			
Observations	43673		43673	
Rsquare	0.55		0.33	

Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table 5: Autoregresive coefficients estimation
Dependent variable: Residuals in OLS (ROLS) regression of NF-income

	OLS Ar(1)	*OLS Ar(2)*	*WG Ar(1)*	*WG Ar(2)*	*GMM Ar(1)*	*GMM Ar(2)*	*GMM Ar(3)*
ROLS(t-1)	0.662	0.515	-0.075	-0.077	0.242	0.325	0.199
	(0.008)**	(0.016)**	(0.013)**	(0.017)**	(0.036)**	(0.062)**	(0.100)*
ROLS(t-2)		0.245		-0.089		0.074	0.007
		(0.016)**		(0.017)**		(0.033)*	(0.052)
ROLS(t-2)							-0.009
							(0.038)
Constant	0.014	0.019	0.030	0.049			
	(0.003)**	(0.004)**	(0.002)**	(0.003)**			
Sargan					33.36	26.33	23.00
(d.f.)					27	25	22
p-value					0.19	0.39	0.40
m1					-12.29	-8.03	-5.27
m2					0.68	-0.86	0.62
Observations	12146	5977	12146	5977	5977	3126	1553
R-squared	0.42	0.48					
Number of individuals			6132	2845	2845	1573	759

Robust standard errors in parentheses. Sargan test, m1 and m2 in two step estimation.
* significant at 5%; ** significant at 1%

Table 6: Estimates of the components of the nonfinancial income process
(in log, euros, prices 1995)

Mean, standard deviation and correlations

a. Mean and standard deviation	*Mean*	*St. Deviation*
Nonfinancial income	9.43	0.52
Effect of age and time trend	0.50	0.10
Time invariant observables characteristics	8.87	0.27
Unobservable individual effect	-0.01	0.24
Idiosyncratic shocks	0.06	0.16

b. Correlations	Nonfinancial income	Effect of age and time trend	Time invariant observables characteristics	Unobservable individual effect	Idiosyncratic shocks
Nonfinancial income	1				
Effect of age and time trend	0.238	1			
Time invariant observables characteristics	0.570	0.118	1		
Unobservable individual effect	0.721	-0.035	0.034	1	
Idiosyncratic shocks	0.583	-0.009	0.038	0.817	1

3061 observations

Table 7: Estimates of a Arch model
Dependent variable: Estimated individual Income Shocks Square (ISS)

	OLS1	*OLS2*	*WG1*	*WG2*	*GMM1*	*GMM2*
ISS(t-1)	0.148	0.153	-0.316	-0.371	-0.032	-0.016
	(0.056)**	(0.074)*	(0.030)**	(0.040)**	(0.042)	(0.066)
ISS(t-2)		0.136		-0.175		0.003
		(0.043)**		(0.036)**		(0.058)
Constant	0.048	0.040	0.075	0.090		
	(0.004)**	(0.005)**	(0.003)**	(0.005)**		
Sargan					6.22	3.66
(d.f.)					9	7
p-value					0.72	0.82
m1					-3.22	-1.38
m2					0.92	-0.47
Observations	1553	794	1553	794	794	338
R-squared	0.03	0.08	0.12	0.21		
Number of individuals			759	456	456	266

Robust standard errors in parentheses. Sargan test, m1 and m2 in two step estimation.
* significant at 5%; ** significant at 1%

Tabla 8: Estimation of the auxiliary (logit) model
Dependent variable: Participation in the risky asset markets

	Education group			
	Elementary	***Secondary***	***High School***	***College and postgraduated***
Age	-0.042	-0.034	-0.065	-0.011
	(0.028)	(0.020)	(0.020)**	(0.039)
Non-financial income (log)	2.321	1.421	1.507	1.757
	(0.288)**	(0.224)**	(0.217)**	(0.440)**
Cash-on-hand/NFI (log)	2.581	2.964	3.499	2.182
	(0.397)**	(0.424)**	(0.397)**	(0.755)**
CoH/NFI square	-5.122	-7.863	-9.674	-5.736
	(1.718)**	(2.224)**	(1.882)**	(3.616)
Family size	-0.169	-0.271	-0.059	-0.173
	(0.120)	(0.088)**	(0.086)	(0.156)
No. labour income earners	-0.456	0.075	-0.243	0.153
	(0.163)**	(0.124)	(0.128)	(0.279)
Working Status	0.084	-0.610	0.063	0.314
	(0.324)	(0.272)*	(0.270)	(0.555)
House Ownership	0.624	0.128	0.234	-0.351
	(0.290)*	(0.197)	(0.209)	(0.461)
Constant	-22.132	-13.174	-14.301	-17.092
	(2.736)**	(2.056)**	(2.047)**	(4.366)**

Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table 9: Generalized Indirect Inference estimation of participation costs (in percentage of non-financial income)

Education group	No. Observations	Estimated Cost	Robust Std. Error	t	95% Conf. Interval	
College and postgraduated	270	0.175	0.06140	2.84998	0.05465	0.29535
High School	964	1.650	0.32542	5.07032	1.01217	2.28783
Secondary	926	3.350	0.34595	9.68355	2.67194	4.02806
Elementary	795	6.000	0.42006	14.28367	5.17668	6.82332

Table 10: Implicit monetary participation costs (in euros 2006)

Education group	No. Observations	Estimated Cost (1)	Average Non-financial Income (2)	Implicit Monetary Costs (2)
College and postgraduated	270	0.175	59076	103
High School	964	1.650	43574	719
Secondary	926	3.350	34550	1157
Elementary	795	6.000	18765	1126

(1) in percent of non-financial income

(2) In euros 2006

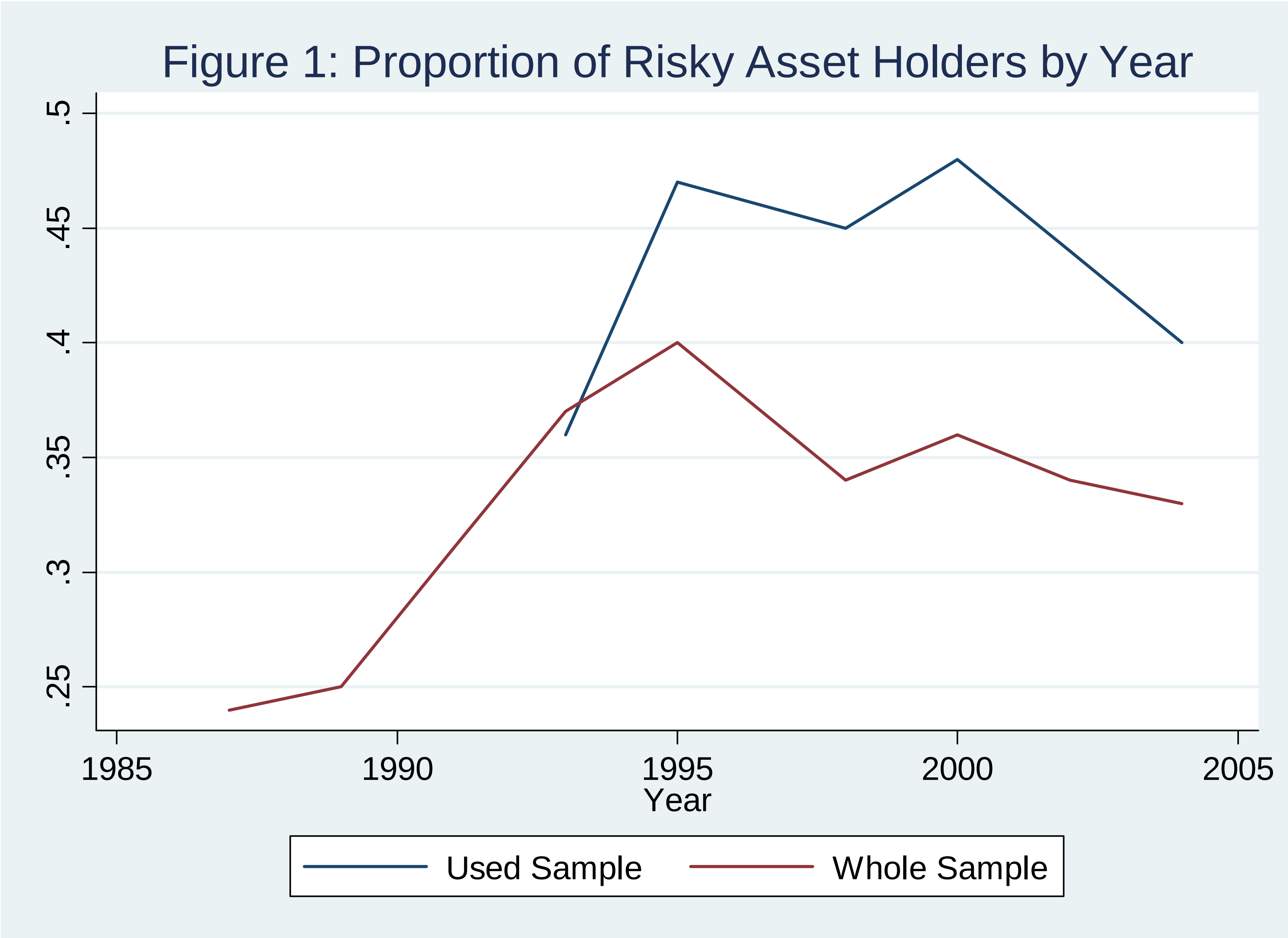


Figure 1: Proportion of Risky Asset Holders by Year
.25
.3
.35
.4
.45
.5
1985
1990
1995
2000
2005
Year
Used Sample
Whole Sample

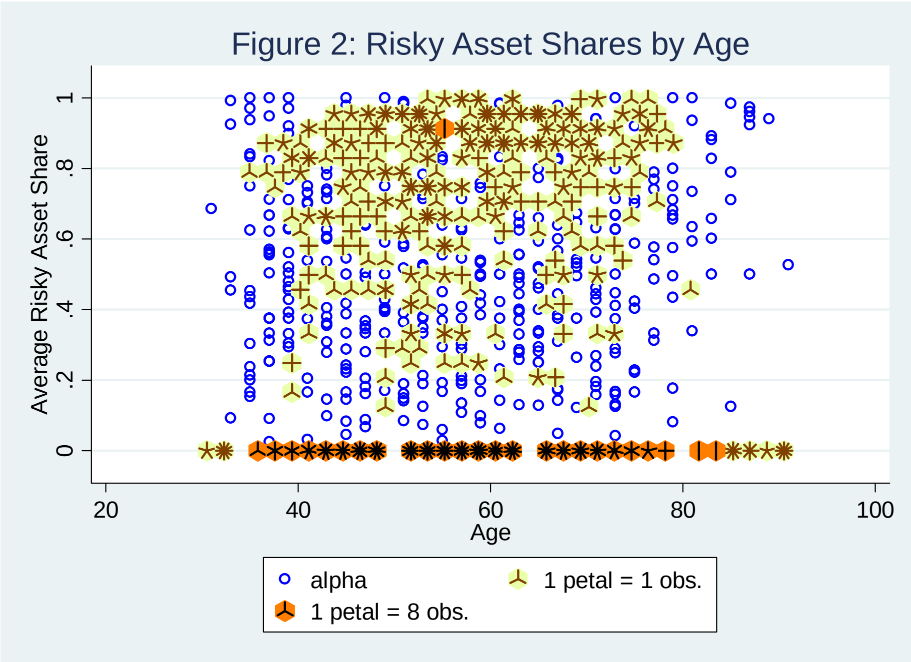


Figure 2: Risky Asset Shares by Age
Average Risky Asset Share
0
.2
.4
.6
.8
1
20
40
60
80
100
Age
alpha
1 petal = 1 obs.
1 petal = 8 obs.

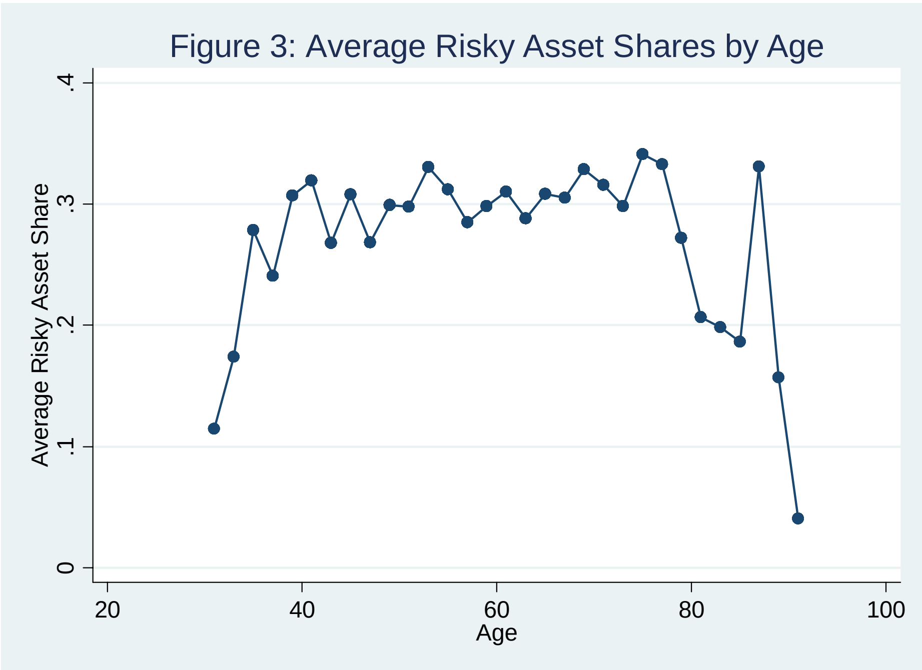


Figure 3: Average Risky Asset Shares by Age
Average Risky Asset Share
0
.1
.2
.3
.4
20
40
60
80
100
Age

Figure 4: Risky Asset Shares by Cash-on-Hand/NF-Income Ratio

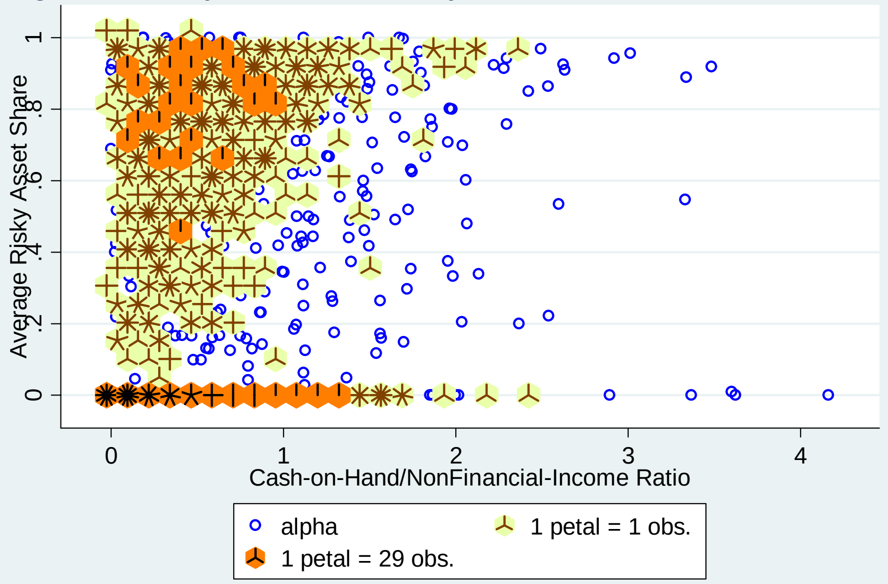
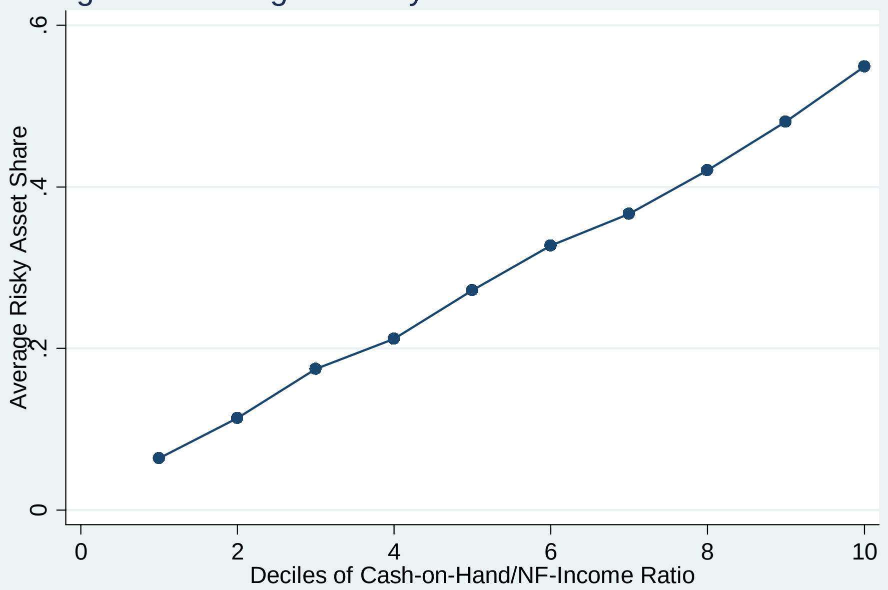


Figure 5: Average RAS by Cash-on-Hand/NF-Income Ratio



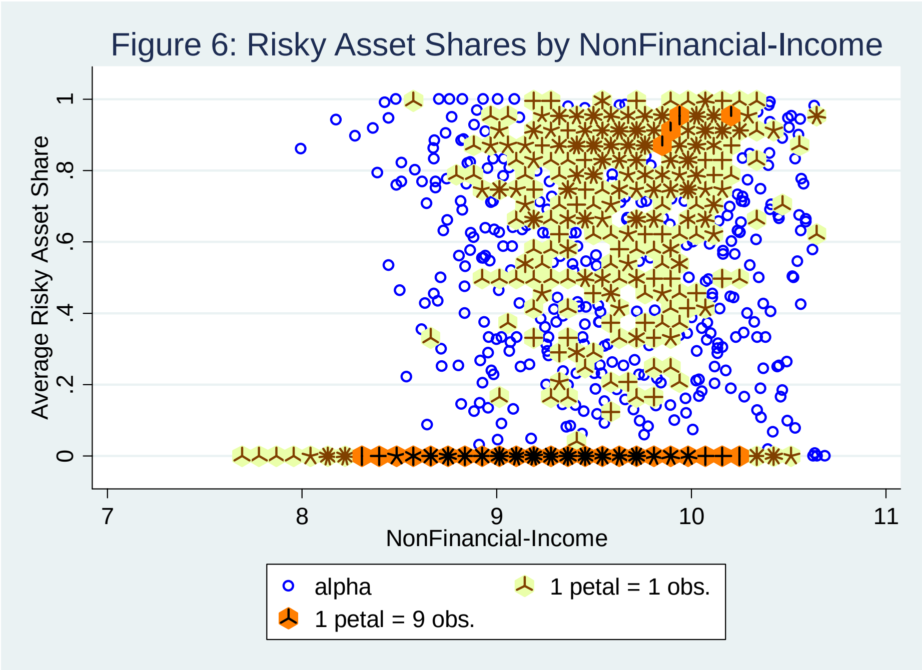


Figure 6: Risky Asset Shares by NonFinancial-Income
Average Risky Asset Share
0
.2
.4
.6
.8
1
7
8
9
10
11
NonFinancial-Income
alpha
1 petal = 1 obs.
1 petal = 9 obs.

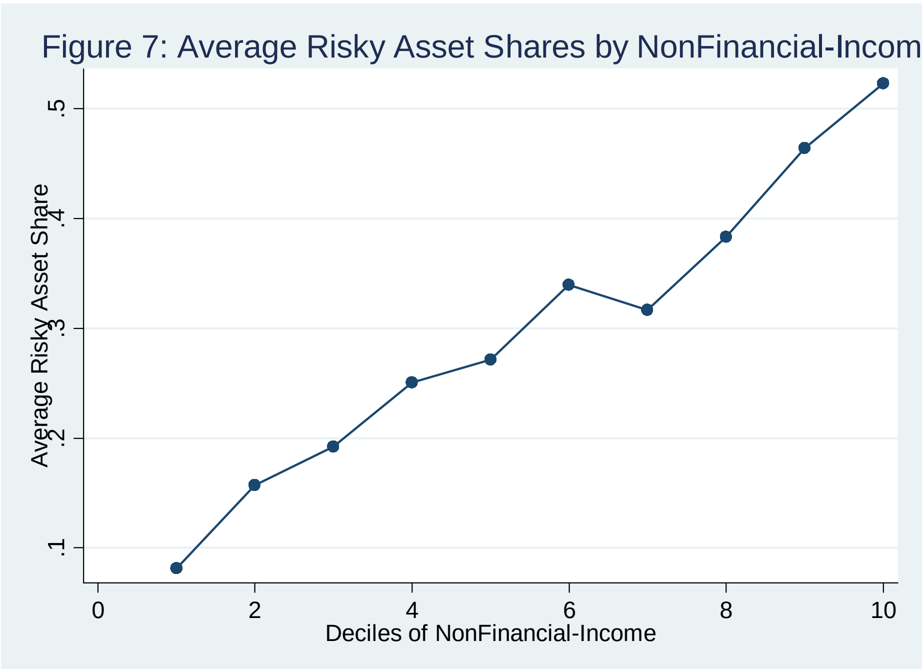


Figure 7: Average Risky Asset Shares by NonFinancial-Income
Average Risky Asset Share
.1
.2
.3
.4
.5
0
2
4
6
8
10
Deciles of NonFinancial-Income

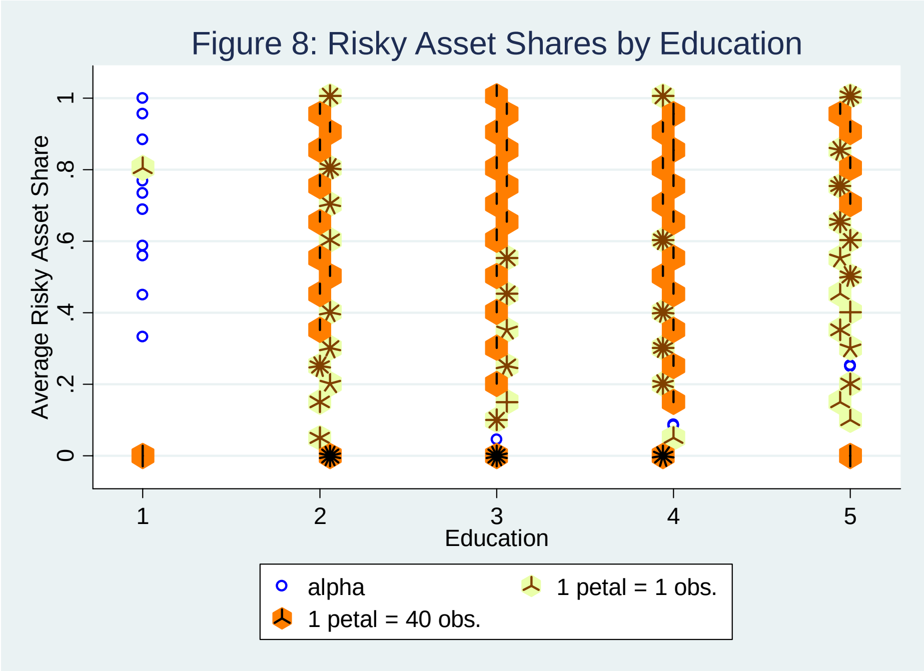


Figure 8: Risky Asset Shares by Education
Average Risky Asset Share
0
.2
.4
.6
.8
1
1
2
3
4
5
Education
alpha
1 petal = 1 obs.
1 petal = 40 obs.

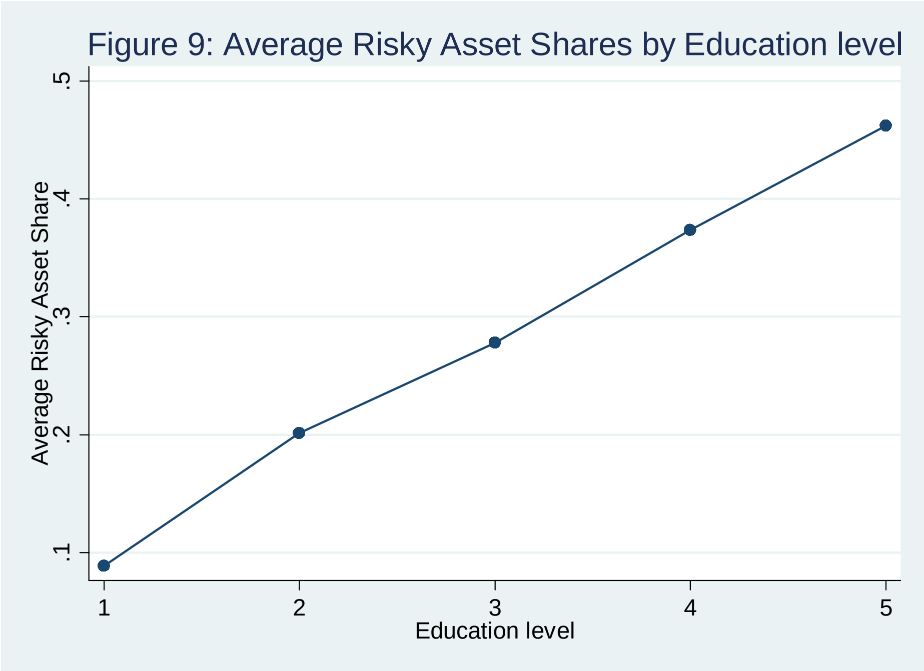


Figure 9: Average Risky Asset Shares by Education level
Average Risky Asset Share
.1
.2
.3
.4
.5
1
2
3
4
5
Education level

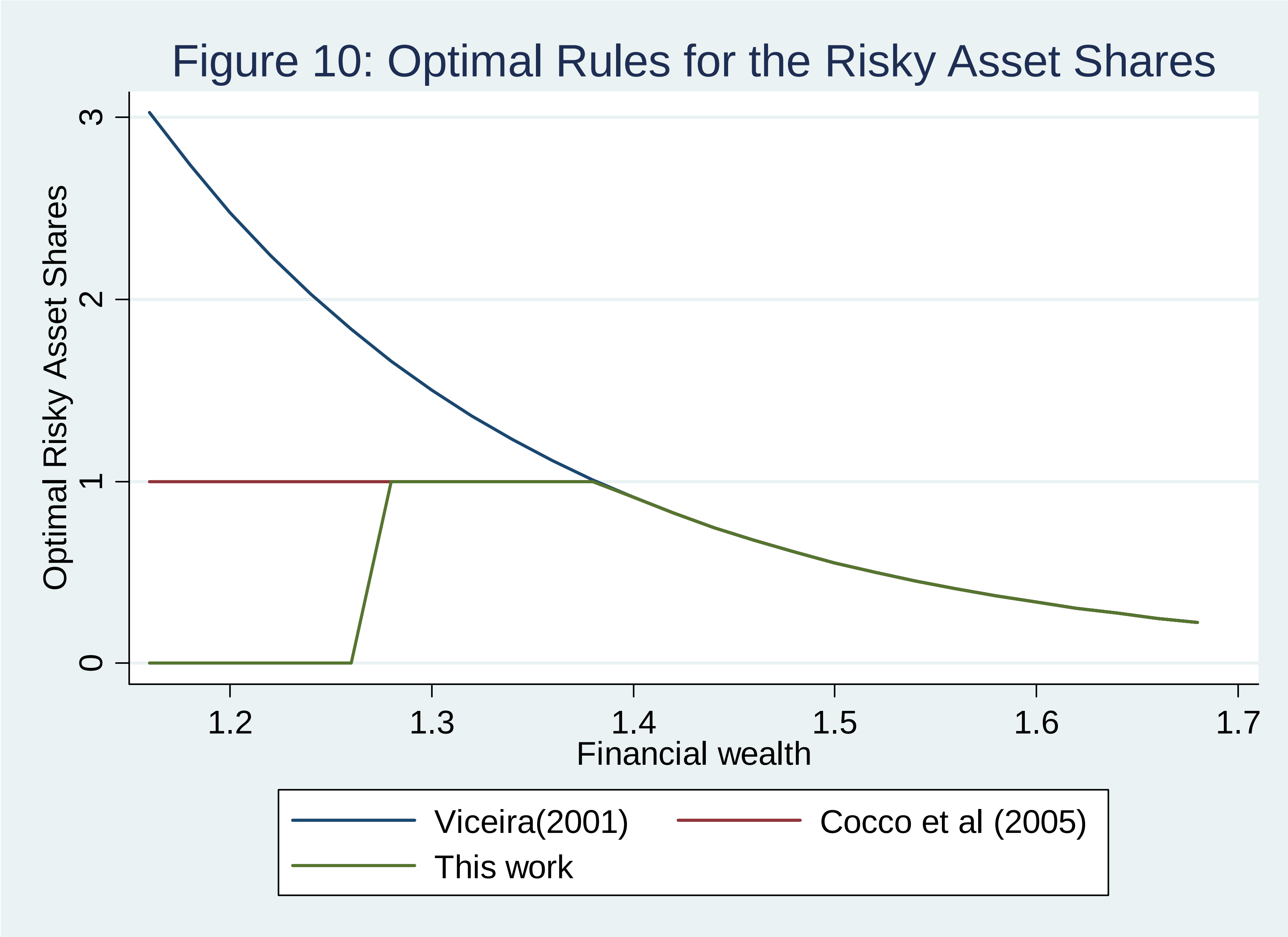


Figure 10: Optimal Rules for the Risky Asset Shares
Optimal Risky Asset Shares
0
1
2
3
1.2
1.3
1.4
1.5
1.6
1.7
Financial wealth
Viceira(2001)
Cocco et al (2005)
This work

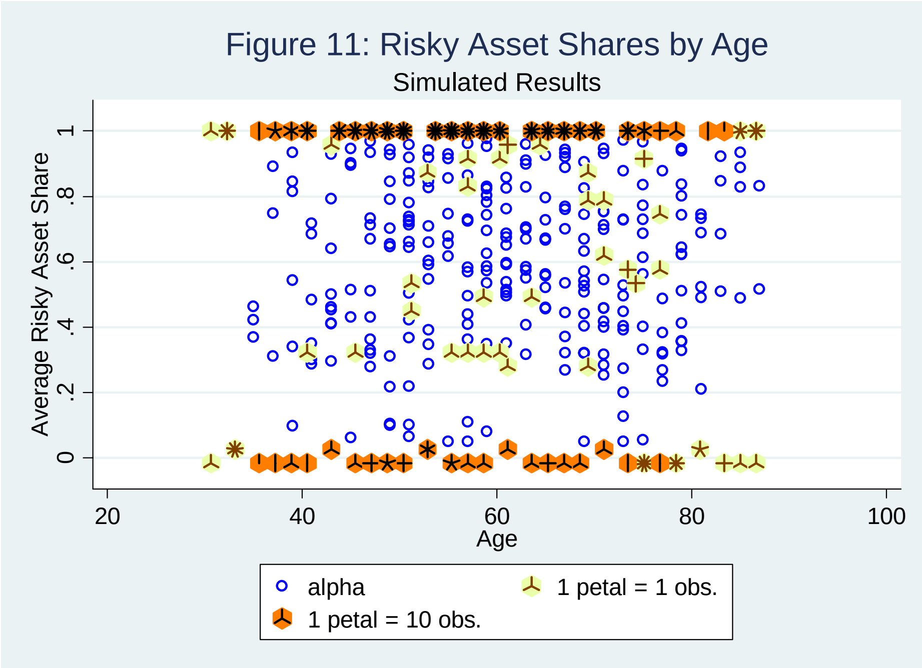


Figure 11: Risky Asset Shares by Age
Simulated Results
Average Risky Asset Share
0
.2
.4
.6
.8
1
20
40
60
80
100
Age
alpha
1 petal = 1 obs.
1 petal = 10 obs.

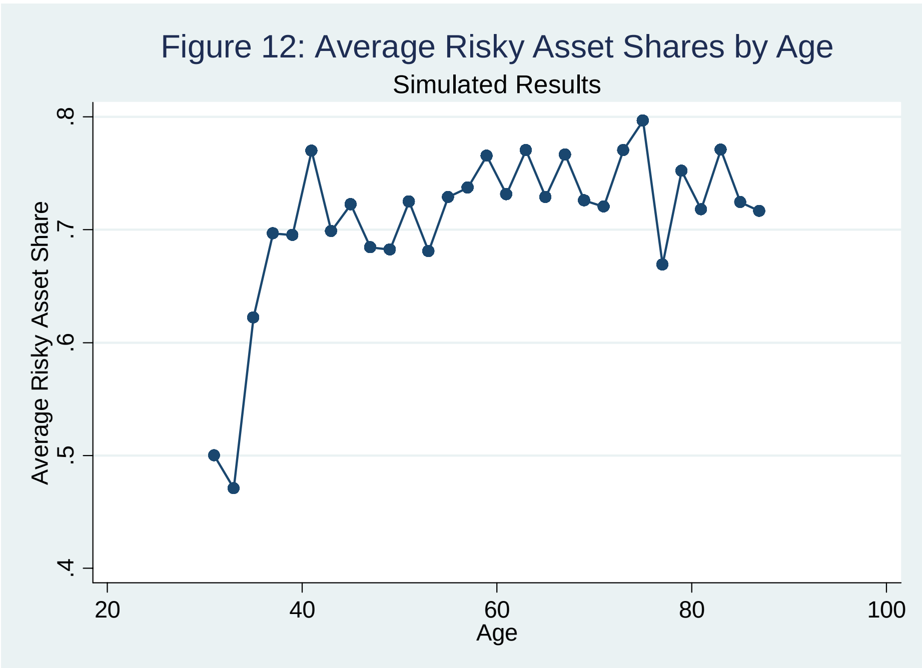


Figure 12: Average Risky Asset Shares by Age
Simulated Results
Average Risky Asset Share
.4
.5
.6
.7
.8
20
40
60
80
100
Age

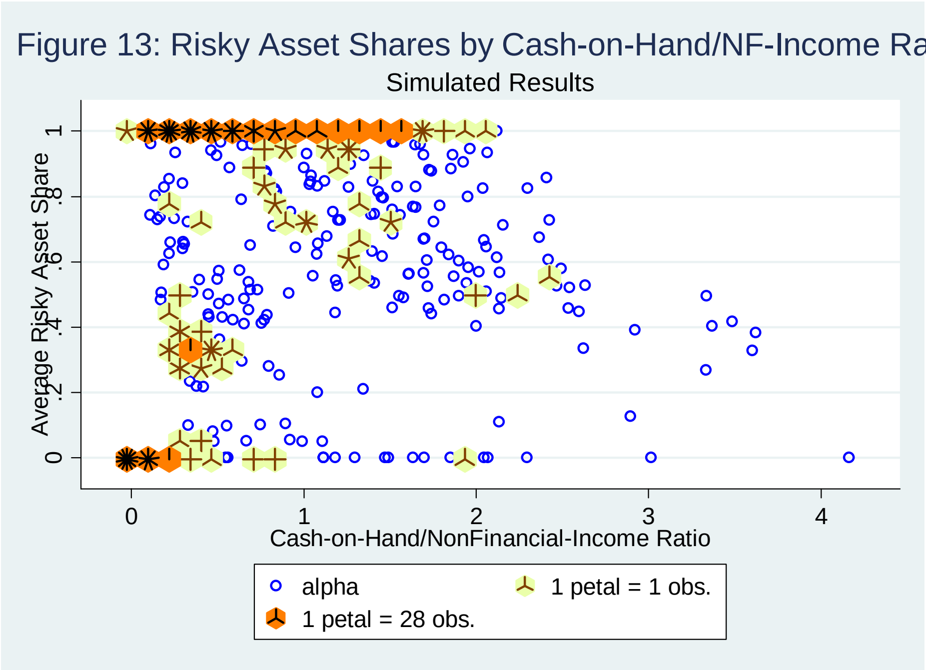


Figure 13: Risky Asset Shares by Cash-on-Hand/NF-Income Ratio
Simulated Results
Average Risky Asset Share
0
.2
.4
.6
.8
1
0
1
2
3
4
Cash-on-Hand/NonFinancial-Income Ratio
alpha
1 petal = 1 obs.
1 petal = 28 obs.

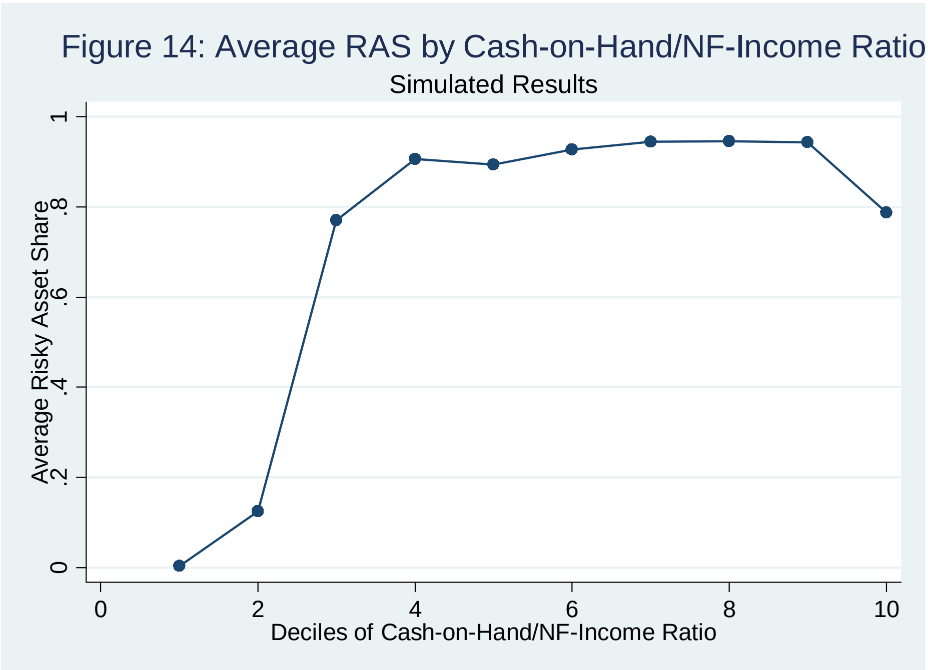


Figure 14: Average RAS by Cash-on-Hand/NF-Income Ratio
Simulated Results
Average Risky Asset Share
0
.2
.4
.6
.8
1
0
2
4
6
8
10
Deciles of Cash-on-Hand/NF-Income Ratio

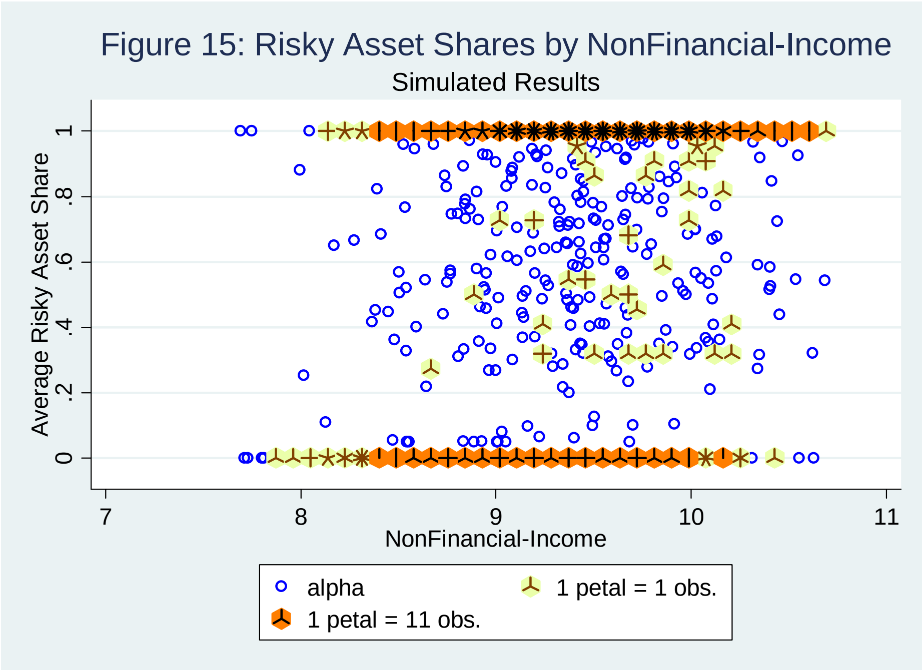


Figure 15: Risky Asset Shares by NonFinancial-Income
Simulated Results
Average Risky Asset Share
0
.2
.4
.6
.8
1
7
8
9
10
11
NonFinancial-Income
alpha
1 petal = 1 obs.
1 petal = 11 obs.

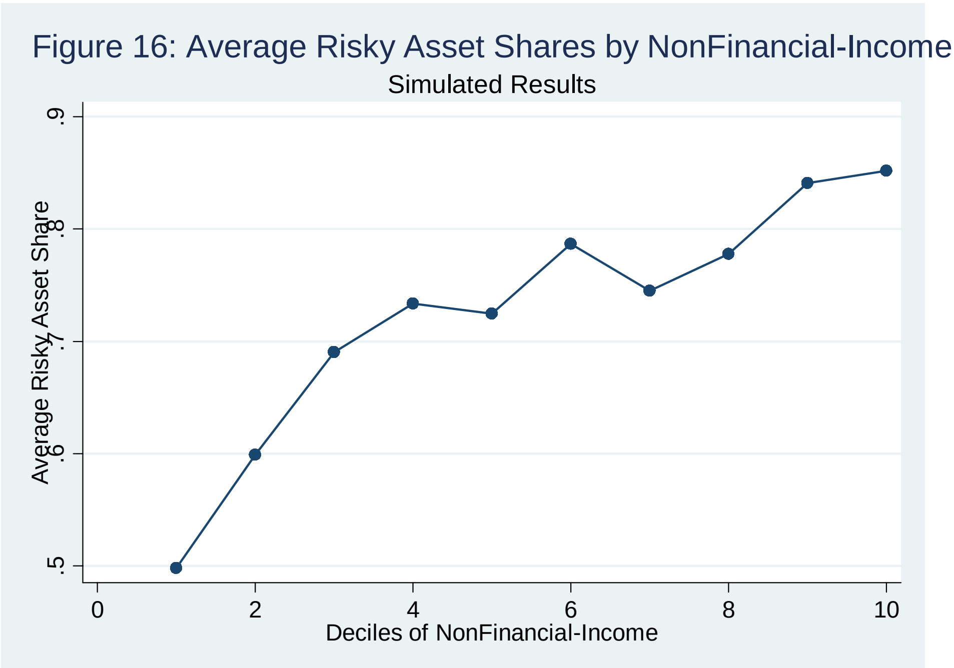


Figure 16: Average Risky Asset Shares by NonFinancial-Income
Simulated Results
Average Risky Asset Share
.5
.6
.7
.8
.9
0
2
4
6
8
10
Deciles of NonFinancial-Income