Bayesian Analysis of the Two-Part Model with Fractional Response: Application to Household Portfolio Choice

Olena Stavrunova*
UTS

Oleg Yerokhin
University of Wollongong

This Draft: September 2008
(Preliminary)

Abstract

This paper studies determinants of household stock marker participation and proportion of risky assets in household financial wealth using recent data from Australia. The methodological novelty of the paper consists in addressing in a systematic fashion the two prominent features of the data: fractional nature of the proportion of financial wealth invested in stocks and prevalence of zeros which stems from the fact that many households do not participate in the stock market. The dependent variable in this case is a mixture of discrete and continuous outcomes, with continuous outcome bounded between zero and one. To study participation and share decisions jointly the paper proposes a two-part model which combines a probit model for participation decision and a linear regression model for the logistic transformation of the fraction of financial wealth invested in stocks. To accommodate possible deviations from normality in the share of risky assets conditional on participation, the transformed share is modelled as having a discrete mixture of normals distribution. The paper then compares posterior distributions of marginal effects of the covariates on participation and share decisions across competing models which include tobit, two-part normal model for untransformed share, two-part normal model for logistically transformed share and two-part mixture of normals model for transformed share. We find that for variables which have a larger explanatory power for participation than for share, the tobit model tends to overpredict marginal effects on share among participants. On the other hand, there seems to be little difference in the marginal effects of covariates on share conditional on participation implied by different versions of the two-part model. Empirical results suggest that education, age, net worth, planning horizon and risk attitudes are the main factors which affect households' exposure to risky assets in Australia.

Keywords: Portfolio Choice, Background Risk, Bayesian Inference, Mixture Models, MCMC
JEL Codes: C11, C24, D14, G11

*Corresponding Author: School of Finance and Economics, University of Technology, Sydney, P.O. Box 123, Broadway, NSW 2007, Australia. Email: olena.stavrunova@uts.edu.au.

1 Introduction

Portfolio choices made over the lifecycle of a household determine the rate of growth of personal wealth and household's standard of living after retirement. Empirical studies of household financial behavior typically document stock market participation rates below 50 percent (Bertaut (1998); Guiso, Haliassos, and Jappelli (2002)) and considerable heterogeneity in the share of wealth invested in stocks among participating households (Heaton and Lucas (2000)). These findings are at odds with the standard portfolio choice theory which predicts that, given the historical equity premium, all households should be willing to invest at least some part of their wealth in the stock market (Merton (1969)), and that the optimal mix of the risky and safe assets in an individual portfolio should be independent of the risk aversion and wealth level (Tobin (1958)). Moreover, estimates based on the calibrated lifecycle models of consumption and portfolio choice imply that welfare loss due to non-participation is equivalent to the 2% reduction in the annual consumption (Cocco et al. (2005)).

Literature on portfolio choice points to fixed participation cost and various sources of non-diversifiable background risk such as business equity and volatile labour income as possible explanations for the low participation rates and cross-sectional variation in the shares of risky assets. Several studies found that foregone earnings of the households which do not hold public equity are relatively small, suggesting that a moderate fixed cost can explain some of the variation in participation rates (Vissing-Jorgensen (2004) (2004), Paliela (2007)). Other studies suggest that households differ in their degree of exposure to the background risk (arising because of the uncertainty about future returns to human capital, business equity, real estate investments and other factors) and that high exposure is associated with lower share of risky assets in household portfolios. For example, Guiso et al. (1996) have shown that subjective expectation of future borrowing constraints and negative income shocks decreases the willingness to hold risky assets among Italian households. Heaton and Lucas (2000b) find that variability of business income reduces the share of risky assets in total wealth among business owners, and that exposure to the employer stock reduces the share of other risky assets for non-entrepreneurs. Hochguertel (2003) documents that Dutch households with higher uncertainty about future labour income tend to tilt their portfolios towards safe assets.

This paper studies the determinants of stock market participation and exposure to risky assets (public equity) conditional on participation using data from a representative survey of Australian households. The active privatization policy of the 1990s and introduction of the system of mandatory retirement contributions (Superannuation Guarantee) have significantly expanded the ranks of Australian shareholders in the last two decades, resulting in one of the highest stock market participation rates in the world (Giannetti and Koskinen (2007)). At the same time, there has been little systematic analysis of the determinants of the stock market participation in Australia. In this paper we use data from the Household Income and Labour Dynamics in Australia (HILDA) survey to construct household level measures of planning horizon, risk attitudes and other characteristics and study their impact on the structure of household portfolios.

Similar to studies of household portfolios in other countries, we document strong impacts of education, age, risk attitudes and net worth levels on the decision to hold public equity. We find moderate effect of education and strong effect of risk attitudes on the proportion of wealth invested in shares among those who participate in the stock market. Moreover, the results suggest that controlling for wealth level and demographic characteristics, households in which head is self-employed are less likely to participate in the stock market and, conditional on participation, tend to invest a smaller share of their liquid financial wealth in the stocks.

The methodological novelty of the paper consists in addressing in a systematic fashion the two prominent features of the data: fractional nature of the proportion of financial wealth invested in stocks and prevalence of zeros which stems from the fact that many households do not participate in the stock market. The dependent variable in this case is a mixture of discrete and continuous outcomes, with continuous outcome bounded between zero and one.

Several approaches have been used in the literature to model stock market participation and share of wealth invested in stocks jointly. The most popular approach is the tobit model, which allows the response to be bounded between zero and one while simultaneously accommodating clustering at zero values. The main disadvantage of this model is that it assumes that the participation and share decisions are affected by the same covariates and that the signs of marginal effects of these covariates on participation and share decision are the same. Another approach is to employ a probit model to analyse participation decision and a linear regression model to analyse determinants of share of risky assets among participants. This approach adds more flexibility compared to the tobit model by allowing for differences in marginal effects of covariates on participation and allocation decisions. However, this approach does not restrict the conditional expectation of share to be between zero and one and does not take into account conditional heteroscedasticity and nonlinear effects of covariates which result when the dependent variable is bounded.

To allow for different effects of covariates on participation and share decisions and to address the bounded nature of the share variable we propose a two part model which combines a probit model for participation decision and a linear regression model for the logistic transformation of the fraction of financial wealth invested in stocks. To accommodate possible deviations from normality in the transformed share conditional on participation, the transformed share is modelled as having a discrete mixture of normals distribution.

The paper takes Bayesian approach to inference and employs Markov Chain Monte Carlo simulation algorithms to access the joint posterior distributions of parameters of the competing models which include tobit, two-part normal model for untransformed share, two-part normal model for logistically transformed share and two-part mixture of normals model for transformed share. We use the method proposed by Chib (2005) to evaluate marginal likelihoods of the competing models.

This selection procedure favors the two-part normal model for logistically transformed share over other alternatives. The paper then compares marginal effects of the covariates on participation and share decisions across the four models. We find that most of the covariates have a larger explanatory power for participation than for share variable, and that because of confounding of the two effects inherent in the tobit model, it tends to overpredict the effects of covariates on share among the participants. On the other hand, there seems to be little difference in the marginal effects of covariates on share conditional on participation implied by different versions of the two-part model.

The rest of the paper is organized as follows. The data used in the paper is described in the next section. Section 3 presents two-part models with normal and mixture of normals disturbances as well as the tobit model and develops an MCMC algorithms for the Bayesian inference in these models. Section 4 derives expressions for marginal effects. Section 5 discusses the empirical results and section 6 concludes.

2 Data and Sample Construction

The data used in this paper comes from the Household Income and Labour Dynamics in Australia (HILDA) survey (Wooden and Watson (2007)). HILDA is a nationally representative longitudinal survey of Australian households. This paper uses data from the second wave of HILDA administered in 2002, which contains a wealth module with detailed information on the composition of household's asset and liabilities in that year. In total, wave 2 of HILDA contains data on 7245 households. We restrict our sample to single-family households which do not include other related or unrelated members, except children. For couple households we define household head as male head of household.

The data at our disposal provides information on the value of public equity held by household either directly or through a mutual fund, which we take as our measure of risky asset holdings. Financial wealth is defined as a sum of bank accounts, cash investments, public equity investments, trust funds and life insurance. The two dependent variables in our analysis are the binary indicator of stock market participation and the share of public equity in financial wealth.

Asset variables which represent background risk faced by a household or alternative investment options include housing equity, business equity and retirement wealth (superannuation). Under Australian mandatory superannuation system employers must contribute a fixed percentage of employee's wage into an account held with a superannuation fund. Superannuation funds typically offer a range of investment options differing in their degree of exposure to stock market risk. As a result superannuation assets can potentially serve as a substitute to direct stock market investment. Because our dataset does not contain information which could be used to assess background risk of household's superannuation wealth, theory does not restrict the sign of the effect of superannuation on stock market participation and the share of financial wealth invested in risky assets. Also, because superannuation assets are accessible upon retirement and only after the age of 55, in our

Table 1: Variable definition and summary statistics (number of observations: 4078)

Variable	Definition	Mean	SD
Financial variables			
equity	=1 if holds public equity	0.47	0.49
fw	Financial Wealth/10000	63.4	204.6
share	share of risky assets in financial wealth among participants	0.55	0.33
nw	(Total assets - total debts)/10000	48.9	68.6
income	Household Income/10000	4.7	5.0
super w	Super, not retired/10000	7.7	14.5
super r	Super, retired/10000	1.9	10.3
bizeq	Own business equity/10000	5.0	32.33
housingeq	Housing equity/10000	18.9	23.9
Household Characteristics			
age	age of household head (HH)	49.8	15.2
edub	=1 if HH has bachelor qualification	0.24	0.42
edud	=1 if HH has advanced diploma	0.37	0.48
edus	=1 if HH is high school graduate	0.09	0.29
olf	=1 if HH is out of labor force	0.28	0.45
unemployed	=1 if HH is unemployed	0.02	0.13
self-employed	HH self-empl.	.10	.31
youngchild	n. of children under 0-15 yrs.	0.53	0.94
health	HH has poor self-assessed health	.17	.37
horizon1	planning horizon next 2-4 yrs.	0.13	0.34
horizon2	planning horizon next 5-10 yrs.	0.25	0.43
risk	HH willing to take high risks	0.12	0.33
nesb	HH has non-English speak. bkgd.	0.12	0.32

empirical specification we distinguish between retiree's and non-retiree's superannuation. Ideally one would want to treat retiree's superannuation as a part of household financial wealth, but because risk profile of household's superannuation assets is not observable it is not possible to classify them as either safe or risky. Therefore, we choose to treat superannuation as another type of asset which can affect portfolio choice. To represent background risk stemming from uncertainty about future labour income we construct binary variable **selfemployed**, which is equal to one if household head is self-employed and is equal to zero otherwise. In the previous studies (e.g. Guiso et al. (1996) and Hochguertel (2003)) measures of subjective labour income uncertainty were found to have statistically significant but relatively small effects on participation and share of risky assets in financial portfolios.

Following other studies of household portfolios we model decisions whether to invest in risky assets and how much to invest conditional on the set of household's socio-economic characteristics, net worth, income, housing equity, health status, risk attitudes, planning horizon and background risk. Household socio-economic characteristics include age and education of household head, number of young children living in household and labour force status of household head (working, unemployed, out of labour force). Household net worth is defined as the difference between total household assets and liabilities. Attitude towards risk is measured using household head's response to the question in which he/she were asked to select a statement best describing the amount of financial risk that he/she were willing to take with spare cash (cash used for savings or investment) from one of the following statements:

1. *I take substantial financial risks expecting to earn substantial returns;*
2. *I take above-average financial risks expecting to earn above-average returns;*
3. *I take average financial risks expecting to earn average returns;*
4. *I am not willing to take any financial risks;*
5. *I never have any spare cash.*

We define risk-loving behavior indicator **risk** as a binary variable, which is equal to one if household head selected statements 1 or 2 and is equal to zero otherwise. Finally, household-level measure of financial planning horizon is constructed from household head's response to the question "*In planning your saving and spending, which of the following time periods is most important to you ?*"

1. *The next week;*
2. *The next few months;*
3. *The next year;*

4. *The next 2 to 4 years;*

5. *The next 5 to 10 years;*

6. *More than 10 years ahead.*

We construct two indicators of the length of financial planning horizon: **horizon1**, which is equal to one if household head selected item 4 and is equal to zero otherwise, and **horizon2**, which is equal to one if household head selected items 5-6 and is equal to zero otherwise. Thus, financial planning horizon enters the model as a set of two dummy variables with planning horizon of less than two years being the omitted category. We further include indicators for non-English speaking background, urban and state residence as predictors of the stock market participation, but not of the share of risky asset in household portfolios in all versions of a two-part model. The assumption here is that these variables are likely to influence amount of information about the stock market available to the household and hence the magnitude of the fixed cost of participating in the market for risky assets (Campbell (2006)). At the same time these variables can be expected to have little influence, conditional on other controls, on the share of wealth invested in stocks. However, because the tobit model assumes that the same set of covariates affects both, participation and share decisions, we include indicators for non-English speaking background, urban and state residence among the covariates which affect both decision.

In the empirical implementation we will use flexible functional forms to accommodate potential non-linear effects of age, net worth and income on participation and allocation decisions. In particular, we introduce second degree polynomials in age and income and a third degree polynomial in net worth into share and participation equations. After eliminating households for which data for at least one variable used in the analysis is missing the final sample consists of 4078 households, of which 1943 or 47% hold some part of their wealth in risky assets. Variable definitions and summary statistics are given in Table 1.

3 Model Specification

This section defines the two-part models with normal and mixture of normals disturbances and the tobit model which will be used to study portfolio choices of Australian households. It also outlines the Markov Chain Monte Carlo algorithms which are used to conduct inference in these models.

3.1 Normal and Mixture of Normals Two-Part Models

Let I_i^* denote the latent utility that an individual derives from stock market participation, let S_i^* denote the logistic transformation of the *potential* proportion of wealth she is willing to invest in the stock market s_i^*:

$$S_i^* = log(\frac{s_i^*}{1 - s_i^*}).$$

The logistic transformation is frequently used in economics and statistics to model fractional response data defined on the interval (0,1) (Kieschnick and McCullough (2003)). The model for the latent vector $[I_i^*, S_i^*]'$ is specified as follows:

$$S_i^* = \boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i} \qquad (1)$$

$$I_i^* = \boldsymbol{\beta}'\mathbf{z}_i + \varepsilon_{2i} \qquad (2)$$

In (1) and (2) the vector of covariates $\mathbf{z}_i$ includes $\mathbf{x}_i$ as well as covariates that belong to the participation equation (2) only. In the normal two-part model the disturbances $\boldsymbol{\varepsilon}_i = [\varepsilon_{1i}, \varepsilon_{2i}]'$ are independently normally distributed:

$$\varepsilon_{1i} \sim N(0, \sigma_1)$$
$$\varepsilon_{2i} \sim N(0, 1).$$

In the two-part model with mixture of normals disturbances the vector ε_{1i} follows a discrete mixture of normal distributions:

$$f(\varepsilon_{1i}|\boldsymbol{\theta}) = \sum_{j=1}^{m} \pi_j \phi(\varepsilon_{1i}|\mu_j, \sigma_{1j}),$$

where $\boldsymbol{\theta}$ denotes the vector of parameters, $\phi(.|a, B)$ denotes probability density function of a normal distribution with mean a and variance B, π_j denotes the probability of mixture component j, m denotes the number of components in the mixture, $\Sigma_{j=1}^m \pi_j = 1$. In this setup mixture components have no structural interpretation because component labels are not identified without prior restrictions. This however is not a concern here because we are using mixture model as a convenient way to relax the normality assumption (Geweke (2007)) and focus only on the permutation invariant functions of interest, such as marginal effects.

Let I_i denote the binary variable which is equal to one if individual i participates in the stock market, and is equal to zero otherwise and assume that

$$I_i = \iota(I_i^* > 0) \qquad (3)$$

where $\iota(a)$ is an indicator function which is equal to one if a is true and is equal to zero otherwise. The potential proportion of wealth invested in stocks, s_i^* and the transformed value S_i^* are only observed when an individual actually participates in the stock market, i.e. when $I_i = 1$. Let s_i^o denote the actual proportion of wealth invested in the stock market, i.e. $s_i^o = s_i^*$ if $I_i = 1$ and $s_i^o = 0$ if $I_i = 0$. Define transformation S_i of the observed share as follows:

$$S_i = log(\frac{s_i^o}{1 - s_i^o}) \cdot \iota(I_i = 1). \qquad (4)$$

The observed data in the two-part model for logistic transformation of share is the vector $[I_i, S_i]'$ for $i = 1, ..., n$, where n denotes sample size. The likelihood function of the two-part model with normal disturbances $L(\boldsymbol{\theta}|\mathbf{Data}, N)$ can be written as:

$$L(\boldsymbol{\theta}|\mathbf{Data}, N) = \prod_{i=1}^{n} \left(1 - \Phi\left(\boldsymbol{\beta}'\mathbf{z}_i\right)\right)^{1-I_i} \cdot \left\langle \Phi\left(\boldsymbol{\beta}'\mathbf{z}_i\right) \cdot \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(S_i - \boldsymbol{\alpha}'\mathbf{x}_i)^2}{2\sigma_1^2}\right)\right\rangle^{I_i}, \quad (5)$$

and that of a two-part model with mixture of normals disturbances $L(\boldsymbol{\theta}|\mathbf{Data}, M)$ can be expressed as follows:

$$L(\boldsymbol{\theta}|\mathbf{Data}, M) = \Pi_{i=1}^{n} \left(1 - \Phi\left(\boldsymbol{\beta}'\mathbf{z}_i\right)\right)^{1-I_i} \cdot \left(\Phi\left(\boldsymbol{\beta}'\mathbf{z}_i\right) \cdot \Sigma_{j=1}^{m} \pi_j \frac{1}{\sqrt{2\pi}\sigma_{1j}} \exp\left(-\frac{(S_i - \mu_j - \boldsymbol{\alpha}'\mathbf{x}_i)^2}{2\sigma_{1j}^2}\right)\right)^{I_i},$$

where $\Phi(a)$ denotes standard normal cdf evaluated at a. [1]

The Bayesian inference in the two-part model with normal disturbances is facilitated by augmenting the vector of data $[S_i, I_i]'$ by the latent utility of stock market participation I_i^*, for $i = 1, ..., n$. The joint probability density function of the augmented data $\mathbf{I}^* = [I_1^*, ..., I_N^*]'$, $\mathbf{S} = [S_1, ..., S_n]'$ and $\mathbf{I} = [I_1, ..., I_n]'$ conditional on the exogenous variables $\mathbf{Z} = [\mathbf{z}_1', ..., \mathbf{z}_n']$ and the vector of parameters $\boldsymbol{\theta}$ can be written:

$$\begin{aligned} & P(\mathbf{I}^*, \mathbf{I}, \mathbf{S}|\mathbf{Z}, \boldsymbol{\theta}) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-(\mathbf{I}^* - \mathbf{Z}\boldsymbol{\beta})'(\mathbf{I}^* - \mathbf{Z}\boldsymbol{\beta})/2\right) \\ \cdot \quad & \Pi_{i=1}^{n} \iota(I_i = 1)\iota(I_i^* \geq 0) + \iota(I_i = 0)\iota(I_i^* < 0)) \\ \cdot \quad & \left(\frac{1}{\sqrt{2\pi\sigma_1}}\right)^{n_{I=1}} \exp\left(-(\mathbf{S}_+ - \mathbf{X}_+\boldsymbol{\alpha})'(\mathbf{S}_+ - \mathbf{X}_+\boldsymbol{\alpha})/2\sigma_1\right) \end{aligned} \quad (6)$$

where $n_{I=1} = \sum_{i=1}^{n} \iota(I_i = 1)$, $\mathbf{S}_+$ is a vector of S_i for observations with $I_i = 1$ and $\mathbf{X}_+$ is a matrix of covariates $\mathbf{x}_i'$ for observations with $I_i = 1$.

The collection of parameters $\boldsymbol{\theta}$ in the normal two-part model consists of σ_1, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. We specify normal prior distributions for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and gamma prior distribution for $h = 1/\sigma_1$, and assume that $\boldsymbol{\alpha}$,$\boldsymbol{\beta}$ and h are independent in the prior:

$$P(\boldsymbol{\theta}) = P(\boldsymbol{\alpha}) \cdot P(\boldsymbol{\beta}) \cdot P(h), \quad (7)$$

where

- $\boldsymbol{\alpha} \sim N(\underline{\boldsymbol{\alpha}}, \underline{H}_\alpha)$
- $\boldsymbol{\beta} \sim N(\underline{\boldsymbol{\beta}}, \underline{H}_\beta)$
- $h \sim gamma(\underline{S}, \nu)$.

[1] For comparison purposes we also estimate a two-part normal model for untransformed share. This model is specified as(1)-(5) with S_i^* and S_i replaced by s_i^* and s_i^o.

The joint posterior distribution of $\boldsymbol{\theta}$ and latent data is proportional to the product of (6) and (7). To approximate this posterior distribution we construct a Gibbs sampling algorithm which iterates between the conditional posterior distributions of I_i^*, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and h. The details of the algorithm are given in the Appendix.

Similar to the normal model, Bayesian inference in the two-part model with mixture of normals disturbances can be conducted by augmenting the observable vector $[S_i, I_i]'$ by the latent utility of stock market participation I_i^* and the latent indicator of mixture component s_i. The latent indicator of mixture component s_i takes on one of the values $1, ..., m$, and $P(s_i = j|\mathbf{z}_i, \boldsymbol{\theta}) = \pi_j$ for $j = 1, ..., m$. The distribution of the disturbances ε_{1i} conditional on the latent indicator of mixture component s_i is normal:

$$\varepsilon_{1i}|(s_i = j, \boldsymbol{\theta}) \sim N(\mu_j, \sigma_{1j}).$$

The following notation will be useful for the presentation of the posterior simulation algorithm in the two-part model with mixture of normals disturbances. Define $\mathbf{C} = [\mathbf{c}_1, ..., \mathbf{c}_{n_{I=1}}]' = [\iota(s_i = j)]$, so that the j^{th} row of the $m \times 1$ vector $\mathbf{c}_i$ is equal to one if $s_i = j$ and is equal to zero otherwise. Also, define $\mathbf{W}_+ = [\mathbf{C}, \mathbf{X}_+]$ and $\boldsymbol{\gamma} = [\mu_1, ..., \mu_m, \boldsymbol{\alpha}']'$. Then the joint probability density function of the augmented data $\mathbf{s} = [s_1, ..., s_n]'$, $\mathbf{I}^* = [I_1^*, ..., I_n^*]'$, $\mathbf{S} = [S_1, ..., S_n]'$ and $\mathbf{I} = [I_1, ..., I_n]'$ conditional on exogenous variables $\mathbf{Z}$ and the vector of parameters $\boldsymbol{\theta}$ can be written:

$$\begin{aligned}
& P(\mathbf{I}^*, \mathbf{I}, \mathbf{S}, \mathbf{s}|\mathbf{Z}, \boldsymbol{\theta}) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-(\mathbf{I}^* - \mathbf{Z}\boldsymbol{\beta})'(\mathbf{I}^* - \mathbf{Z}\boldsymbol{\beta})/2\right) \\
\cdot \quad & \Pi_{i=1}^n (\iota(I_i = 1)\iota(I_i^* \geq 0) + \iota(I_i = 0)\iota(I_i^* < 0)) \qquad (8) \\
\cdot \quad & \sum_{j=1}^m \pi_j \left(\frac{1}{\sqrt{2\pi\sigma_{1j}}}\right)^{n_{I=1}^j} \exp\left(-(\mathbf{S}_+ - \mathbf{W}_+\boldsymbol{\gamma})' Q (\mathbf{S}_+ - \mathbf{W}_+\boldsymbol{\gamma})/2\right)
\end{aligned}$$

where $n_{i=1}^j$ is the number of observations such that $s_i = j$ and $I_i = 1$, and Q is a $n_{I=1} \times n_{I=1}$ diagonal matrix with the diagonal element $q_{ii} = 1/\sigma_{1s_i}$.

The collection of parameters $\boldsymbol{\theta}$ in the two-part model with mixture of normals disturbances consists of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $h_1, ..., h_m$ and $\boldsymbol{\pi}$, where $h_j \equiv \sigma_{1j}^{-1}$ for $j = 1, ..., m$. We specify normal prior distributions for $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, gamma prior distributions for $h_1, ..., h_m$, Dirichlet prior distribution for $\boldsymbol{\pi}$ and assume that $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $h_1, ..., h_m$ and $\boldsymbol{\pi}$ are independent in the prior:

$$P(\boldsymbol{\theta}) = P(\boldsymbol{\beta})P(\boldsymbol{\gamma}) \cdot \prod_{j=1}^m P(h_j) \cdot P(\boldsymbol{\pi}), \qquad (9)$$

where

- $\boldsymbol{\beta} \sim N(\underline{\boldsymbol{\beta}}, \underline{H}_\beta)$
- $\boldsymbol{\gamma} \sim N(\underline{\boldsymbol{\gamma}}, \underline{H}_\gamma)$

- $h_j \sim gamma(\underline{S}_j, \nu_j)$ for $j = 1, ..., m$
- $\boldsymbol{\pi} \sim Dirichlet(r)$

The joint posterior distribution of $\boldsymbol{\theta}$ and latent data in the two-part model with mixture of normals disturbances is proportional to the product of (8) and (9). To approximate this posterior distribution we construct a Gibbs sampling algorithm which iterates between the conditional posterior distributions of I_i^*, s_i, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, h_j and $\boldsymbol{\pi}$. The details of the algorithm are given in the Appendix. The Matlab codes for these two algorithms have passed the joint distribution test suggested in Geweke (2004).

3.2 The Tobit Model

The tobit model assumes that the *potential* proportion of wealth an individual is willing to invest in the stock market s_i^* follows normal distribution:

$$s_i^* = \boldsymbol{\beta}'\mathbf{z}_i + \varepsilon_i, \tag{10}$$

where

$$\varepsilon_i \sim N(0, \sigma), \ i = 1, ..., n.$$

Then the observed proportion of wealth invested in the stock market s_i^o is assumed to be generated by s_i^* as follows:

$$s_i^o = max\{0, s_i^*\}. \tag{11}$$

The likelihood function of the tobit model $L(\boldsymbol{\theta}|\mathbf{Data}, T)$ can be expressed as:

$$L(\boldsymbol{\theta}|\mathbf{Data}, I) = \prod_{i:s_i^o=0} \left(1 - \Phi\left(\boldsymbol{\beta}'\mathbf{z}_i\right)\right) \cdot \prod_{i:s_i^o>0} \cdot \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(s_i^o - \boldsymbol{\beta}'\mathbf{z}_i)^2}{2\sigma}\right).$$

The Bayesian inference in the tobit model is facilitated by augmenting the observables s_i^o by the potential proportion on wealth invested in shares s_i^*, for $i = 1, ..., n$. The joint probability density function of the augmented data $\mathbf{s}^* = [s_1^*, ..., s_n^*]'$, $\mathbf{s} = [s_1, ..., s_n]'$ conditional on the exogenous variables $\mathbf{Z} = [\mathbf{z}_1', ..., \mathbf{z}_n']$ and the vector of parameters $\boldsymbol{\theta}$ can be written:

$$\begin{aligned} P(\mathbf{s}^*, \mathbf{s}|\mathbf{Z}, \boldsymbol{\theta}) &= \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^n \exp\left(-(\mathbf{s}^* - \mathbf{Z}\boldsymbol{\beta})'(\mathbf{s}^* - \mathbf{Z}\boldsymbol{\beta})/2\sigma\right) \\ \cdot \quad & \Pi_{i=1}^n \iota(s_i^o > 0)\iota(s_i^* = s_i^o) + \iota(s_i^o = 0)\iota(s_i^* < 0)). \end{aligned} \tag{12}$$

The collection of parameters $\boldsymbol{\theta}$ in the tobit model consists of σ and $\boldsymbol{\beta}$. We specify normal prior distributions for $\boldsymbol{\beta}$ and gamma prior distribution for $h = 1/\sigma$, and assume that $\boldsymbol{\beta}$ and h are independent in the prior:

$$P(\boldsymbol{\theta}) = P(\boldsymbol{\beta}) \cdot P(h), \tag{13}$$

where

- $\boldsymbol{\beta} \sim N(\underline{\boldsymbol{\beta}}, \underline{H}_\beta)$
- $h \sim gamma(\underline{S}, \nu)$.

The joint posterior distribution of $\boldsymbol{\theta}$ and latent data is proportional to the product of (12) and (13). To approximate this posterior distribution we construct a Gibbs sampling algorithm which iterates between the conditional posterior distributions of s_i^*, $\boldsymbol{\beta}$ and h as suggested in Koop et al. (2007).

4 Marginal Effects

The results of the two-part models and the tobit model can be interpreted by comparing marginal effects of covariates on the outcome variables. For each model we compute *posterior distributions* of the following three sets of marginal effects:

1. The marginal effect of the variable z_{ki} on probability of stock market participation of individual i. For continuous z_{ki} this effect is computed as the derivative of the probability of stock market participation of individual i with respect to z_{ki}:

   $$MEP_{z_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv \frac{\partial Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A)}{\partial z_{ki}}, \tag{14}$$

   where superscript c indicates that the marginal effect $MEP_{z_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}$ is that of a continuous z_{ki}, and A indicates a model for which the effect is computed.

   For discrete z_{ki} the effect is computed as the difference in probabilities of stock market participation of individual i evaluated at adjacent values of z_{ki}:

   $$MEP_{z_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv Prob(I_i^* > 0|\mathbf{z}_{-z_{ki},i}, z_{ki} = a+1, \boldsymbol{\theta}, A) - Prob(I_i^* > 0|\mathbf{z}_{-z_{ki},i}, z_{ki} = a, \boldsymbol{\theta}, A), \tag{15}$$

   where superscript d indicates that the marginal effect $MEP_{z_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}$ is that of a discrete z_{ki}.

2. The marginal effect of the variable x_{ki} on the expectation of the fraction of wealth invested in shares of individual i conditional on participation in the stock market. For continuous x_{ki} this effect is computed as the derivative of the expectation of fraction of wealth invested in stock market of individual i conditional on participation with respect to x_{ki}:

   $$MESC_{x_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv \frac{\partial E(s_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, A)}{\partial x_{ki}}. \tag{16}$$

For discrete x_{ki} the effect is computed as the difference in the expectations of the fraction of wealth invested in shares of individual i conditional on participation evaluated at adjacent values of x_{ki}:

$$MESC_{x_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv E(s_i^*|I_i^* > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a+1, \boldsymbol{\theta}, A) - E(s_i^*|I_i^* > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a, \boldsymbol{\theta}, A). \tag{17}$$

3. The effect of the variable x_{ki} on the unconditional expectation of observed fraction of wealth invested in shares of individual i. This unconditional expectation can be expressed:

$$\begin{aligned} & E(s_i^o|\mathbf{z}_i, \boldsymbol{\theta}, A) = E(s_i^o|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, A) \cdot Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A) \\ + \quad & E(s_i^o|I_i^* \leq 0, \mathbf{z}_i, \boldsymbol{\theta}, A) \cdot Prob(I_i^* \leq 0|\mathbf{z}_i, \boldsymbol{\theta}, A). \end{aligned}$$

Because the observed fraction of wealth invested in shares is zero for individuals who do not participate in the stock market, and the observed fraction s_i^o is equal to the potential fraction s_i^* for individuals who participate in the stock market, the expectation $E(s_i^o|\mathbf{z}_i, \boldsymbol{\theta}, A)$ reduces to:

$$E(s_i^o|\mathbf{z}_i, \boldsymbol{\theta}, A) = E(s_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, A) \cdot Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A). \tag{18}$$

Then the marginal effect of a continuous variable x_{ki} is computed as the derivative of this unconditional expectation with respect to x_{ki}:

$$MESU_{x_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv \frac{\partial E(s_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, A) Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A)}{\partial x_{ki}}. \tag{19}$$

The marginal effect of a discrete x_{ki} is computed as the difference in the unconditional expectations evaluated at adjacent values of x_{ki}:

$$\begin{aligned} & MESU_{x_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, \text{A} \equiv \\ & E(s_i^*|I_i^* > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a+1, \boldsymbol{\theta}, A) Prob(I_i^* > 0|\mathbf{z}_{-x_{ki},i}, x_{ki} = a+1, \boldsymbol{\theta}, A) - \\ & E(s_i^*|I_i^* > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a, \boldsymbol{\theta}, A) Prob(I_i^* > 0|\mathbf{z}_{-x_{ki},i}, x_{ki} = a, \boldsymbol{\theta}, A). \end{aligned} \tag{20}$$

The expressions for the marginal effects (14)-(20) in two-part models with normal and mixture of normals disturbances and in the tobit model are derived in the Appendix. These marginal effects all depend of the vector of covariates $\mathbf{z}_i$, so in general for a given $\boldsymbol{\theta}$ there will be as many marginal effects of the variable z_{ki} as there are individuals in the sample. It has become a standard practice to evaluate marginal effects at sample means or medians of the covariates, and we will follow this convention hereafter. In particular, we evaluate marginal effects for a representative Australian household, which we define as a household whose continuous covariates (household head's age, net worth, income, business equity, housing equity, superannuation) are equal to their sample medians, with no children younger than 16 years old, living in New South Wales and whose household head is employed but not self-employed, is not willing to take risks, has advanced diploma as the highest educational qualification, is in good health, comes from English-speaking background and has planning horizon of less or equal to one year. To obtain the *posterior distributions* of the marginal effects we evaluate model-specific expressions for (14)-(20) for a range of parameters representative of their posterior distribution, i.e. we use draws from the posterior distribution of parameters $p(\boldsymbol{\theta}|\mathbf{Data})$ to approximate the following posterior distributions of marginal effects:

$$p(MEf_{z_{ki}^h}|\mathbf{z}_i = \overline{\mathbf{z}}, \mathbf{Data}, A) = \int p(MEf_{z_{ki}^h}|\mathbf{z}_i = \overline{\mathbf{z}}, \boldsymbol{\theta}, A)p(\boldsymbol{\theta}|\mathbf{Data})d\boldsymbol{\theta}, \tag{21}$$

where $f = \{P, S, SU\}$, $h = \{c, d\}$ and $\overline{\mathbf{z}}$ denotes the vector of covariates $\mathbf{z}_i$ of a representative household. To summarize these posterior distributions, for every marginal effect we report posterior mean, posterior standard deviation and posterior probability that the effect is positive.

5 Empirical results

One of the goals of the paper is to answer two related questions: first, which household characteristics have significant influence on the decision to hold public equity in Australia? Second, what are the main determinants of the share of financial wealth invested in the stock market? To answer these questions we fit four models to the share and participation data: normal and two component mixture two-part models to transformed share data, tobit model to untransformed share and normal two-part model to untransformed share data. For each model the posterior distribution of parameters is obtained from the MCMC chain with 100,000 iterations. First 20,000 iterations are discarded to allow the effect of an initial draw to vanish, and the remaining 80,000 draws are used for the analysis. The hyperparameters of prior distributions of parameters of the four models are presented in the Appendix.

Model selection is based on the comparison of log marginal likelihoods which are computed using the method proposed by Chib (1995). As shown in Table 2, the data favors the normal model on transformed share data over the tobit model, the normal model on untransformed share and the model on transformed share with mixture of normals disturbances. In what follows we treat the

Table 2: Log Marginal Likelihood Comparison

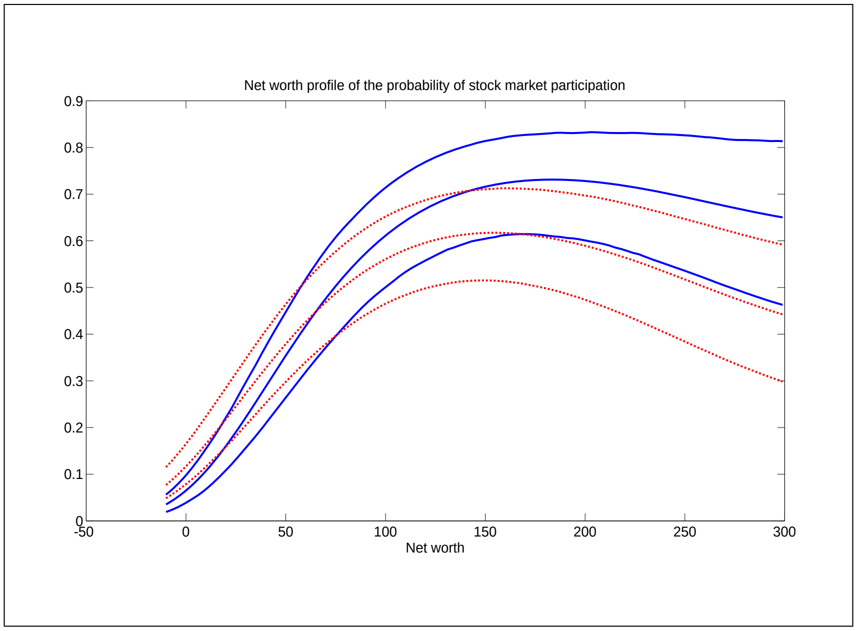
Model	Log of Marginal Likelihood
Tobit	-3126.2
Normal two-part, without transformation	-3028.5
Normal two-part, transformed	**-2445.1**
Mixture of 2 Normals two-part, transformed	-2552

normal model on transformed share as our preferred model and discuss results from this model in detail. We also present results from the tobit model and from the two-part model on transformed share with mixture of normals disturbances and the two-part normal model on untransformed share. As it turns out, all two-part models produce very similar posterior distributions of marginal effects. The empirical results for the normal model on transformed share are presented in tables 3, 5 and 7. Table 3 presents posterior moments of the coefficients and marginal effects in the participation equation. Table 5 contains posterior moments of the coefficients in the share equation and posterior moments of the marginal effect of covariates on the share of risky assets conditional on participation. Posterior moments of the unconditional marginal effects of covariates on the observed share of risky assets are given in Table 7. The results from the tobit model are reported in tables 4, 6 and 8. Table 4 contains coefficients and marginal effects in the participation equation, and tables 6 and 8 present posterior moments of the coefficients and conditional and unconditional marginal effects in the share equation. All marginal effects are evaluated for a representative household as defined in section 4.

As discussed in section 2, we assume that non-English speaking background and state of residence only influence participation decision and therefore can be excluded from the share equation in two-part models. We hypothesize that these variables will have a significant impact on the participation decision because all of them are likely to influence the information about opportunities of investing in the stock market available to a household, while having no influence on the share of risky assets in household portfolios. Consistent with our hypothesis, results presented in Table 3 imply that non-native speaker indicator has a strong negative impact on participation: people with non-English language background are 13 percentage points less likely to invest in stocks. We also find that there exist significant differences in participation rates across states, with residents of Tasmania and Northern Territory (the omitted category) being less likely to hold stocks than those of other states.

As can be seen from Table 3 and Figure 1 households with higher net worth are more likely to invest in stocks. Figure 1 shows the net worth profile of the probability of stock market participation of a representative household implied by tobit (red dashed line) and two-part normal model for transformed share (blue solid line). To construct Figure 1 we evaluated probability of stock market participation as defined in (A-1) for a grid of values of net worth ranging from -50 to 300, with other covariates fixed at values of a representative households, and for a range of parameters representative of their posterior distribution. This way we obtained posterior distributions of the probability of stock market participation for each point of the net worth grid. Figure 1 plots the mean, 5th and 95th percentiles of these distributions implied by the two models for each point of the grid. The net worth profile from the two-part normal model on transformed share (the preferred model) suggests that net worth is a strong predictor of stock market participation, with mean participation probability increasing from 0.05 to 0.70 as net worth increases from \$0 (5th

Figure 1: Net worth profile of the probability to hold stocks; blue solid- two-part model, red dashed-tobit

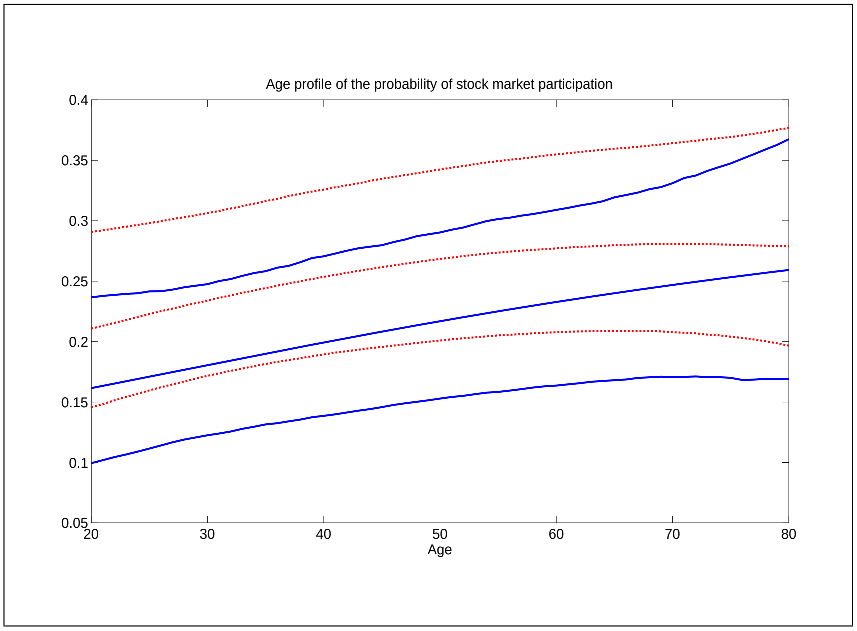


percentile of sample distribution of net worth) to $1,500,000 (95th percentile of sample distribution of net worth). Note, that the tobit model overpredicts the net worth profile for low values of net worth, and underpredicts it for high values of net worth.

Figure 2 shows age profile of the probability to invest in shares. The results from the preferred model imply that the effect of age on participation is small and almost linear: participation probability is increasing by 0.01 per each 10 additional years. The tobit model predicts higher participation for every point of the age grid, but the shape of the age profile implied by the tobit model is similar to that implied by the preferred two-part model.

Table 3 implies that participation is also increasing with the level of education: households headed by persons with 12 years of schooling are 9 percentage points more likely to hold public equity compared to those with less than 12 years of schooling and this result is consistent among the two-part models (see Table 5). Interestingly, additional education beyond 12 years does not seem to increase the probability of participation. Table 3 also implies that planning horizon and attitude towards risk have large positive effects on stock market participation. Effect of risk preferences is especially strong: respondents who report willingness to take high and moderate risks in order to earn a higher return are 18 percentage points more likely to hold risky assets. These results allow us to conclude that risk attitudes is the most significant determinants of the participation decision besides age and wealth. Note that the Tobit model implies smaller effects of planning horizon and risk variables on participation.

Figure 2: Age profile of the probability to hold stocks; blue solid- two-part model, red dashed-tobit

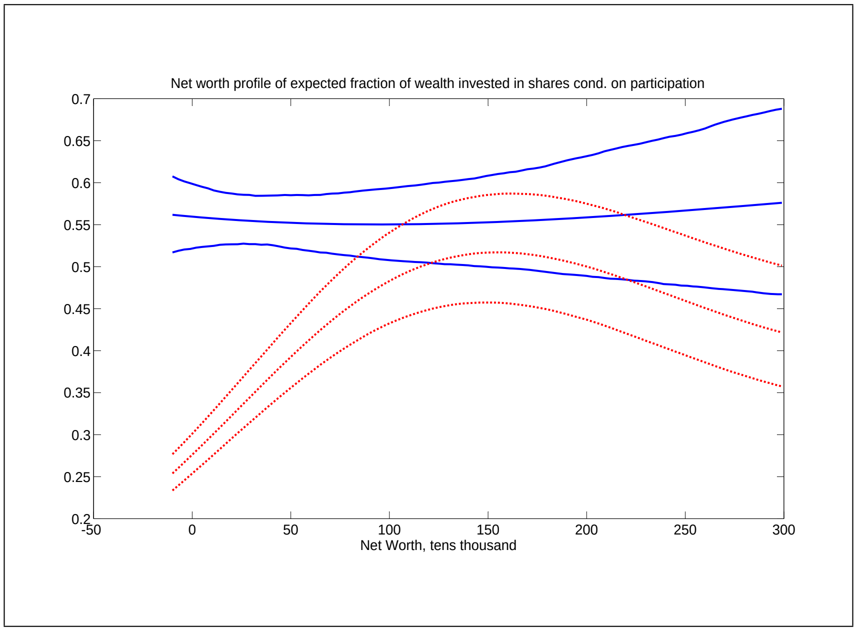


Finally, variables measuring background risk (or alternative investment options) have expected but small effects on stock market participation. For example, households with self-employed heads are 1 percentage point less likely to hold stocks, and households with heads in bad health are 4 percentage points less likely to hold stocks. The effect of business equity is also negative, but rather small. Increasing business equity by $100,000 is associated with 1.4 percentage points decrease in participation probability. We do not find any relationship between stock market participation and superannuation assets: the posterior distributions of marginal effects of superannuation are centered around zero. This implies that households in general do not view their pension wealth as a substitute for direct investments in public equity.

Turning to the share equation we observe that effects of explanatory variables on the share of risky assets conditional on participation are in general quite small. These results are similar to the findings of several studies of household portfolios in other countries (e.g. Vissing-Jorgenson, 2004). Figure 3 demonstrates the net worth profile of share conditional on participation. The preferred two-part model suggests no relationship between wealth and share conditional on participation, while tobit model suggests a hump-shaped profile. This result supports the hypothesis that tobit model confounds effects of covariates on participation and share by attributing strong relationship between participation and wealth present in the data to the relationship between share conditional on participation and wealth .

The age profile of the expected share of wealth invested in stocks conditional on participation

Figure 3: Net worth profile of the fraction of wealth invested in shares conditional on participation; blue solid- two-part model, red dashed-tobit

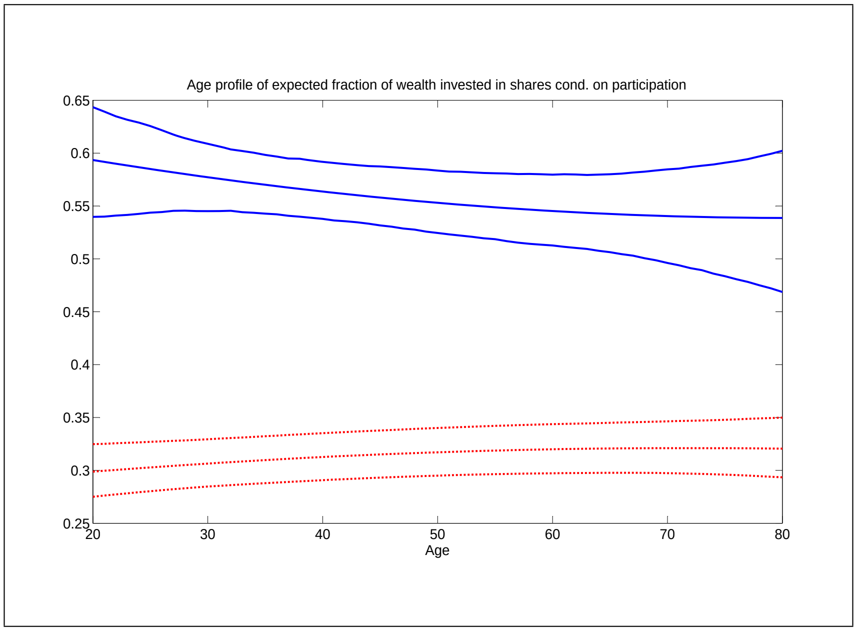


is shown in Figure 4. The figure suggests a weak negative relationship between expected share conditional on participation and age, with share decreasing by about 1 percentage points per each additional 10 years. Interestingly, the tobit model underpredicts share for all points of the age grid and implies a weak positive relationship between age and share.

Education has a positive effect on proportion of financial wealth invested in stocks among those who participate in the stock market: households where household head has a bachelor or higher degree hold 5 percentage points more of their financial wealth in stocks compared to households with less than 12 years of education (see Table 5). Interestingly, the mean marginal effects of education on share in the two-part model are similar to those in the tobit model, although the tobit model suggests that these effects are positive with higher certainty, because the posterior probability that these effects are positive are all equal to 1 in the tobit model, but are less than one in the two-part model (see Table 6).

Most background risk (or alternative investment options) variables have no effect on share conditional on participation: the posterior distributions of marginal effects of business equity and superannuation are all centered around zero. We do find that households with self-employed heads decrease their share of risky assets by 8 percentage points and that households where head has bad self-reported health have about 2 percentage points lower proportion of financial wealth invested in stocks.

Figure 4: Age profile of the fraction of wealth invested in shares conditional on participation; blue solid- two-part model, red dashed-tobit



Turning to effects on expected share (unconditional of participation) we see from figures 5 and 6 that tobit model predicts lower and less steep net worth and age profiles of this expectation compared to the two-part model with logistic transformation. Also, Tables 7 and 8 show that tobit model underpredicts effects of education, risk attitudes and planning horizon on unconditional expectation of fraction of wealth invested in shares compared to the preferred two-part model, and predicts higher probability that marginal effects of business equity and housing equity on this expectation are negative.

Figure 5: Net worth profile of the fraction of wealth invested in shares; blue solid- two-part model, red dashed-tobit

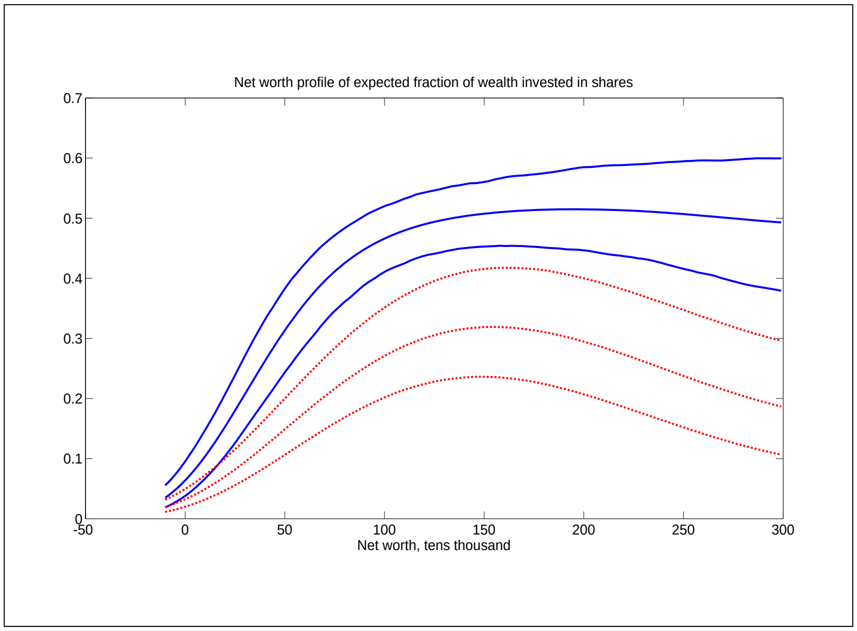
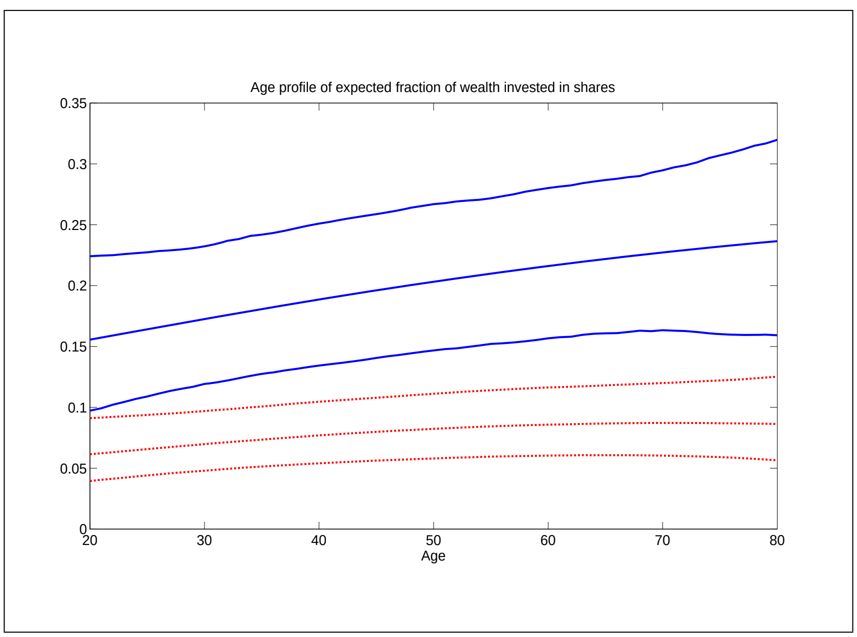


Figure 6: Age profile of the fraction of wealth invested in shares; blue solid- two-part model, red dashed-tobit



6 Conclusion

This paper contributes to the literature on household financial behavior by studying the factors which affect stock market participation and share of risky assets in household portfolios using recent data from Australia. The two decisions are modeled jointly in the framework of a two part model. The joint distribution of error terms is modeled as a finite mixture of normals thus accommodating possible departures from normality. MCMC methods are used to obtain posterior distribution of the parameters and of marginal effects of explanatory variables on the participation and allocation decisions. We find that the data favors normal two-part model with logistically transformed share over other competing models which include normal two-part model for untransformed share, the two-component mixture two-part model for transformed share and the tobit model according to the marginal likelihood criterion.

Results based on the normal two-part model with transformed share imply that education, age, wealth, risk aversion, planning horizon have strong effects on stock market participation. Similar to existing empirical studies of household portfolios in other countries, we find that, conditional on participation, most household characteristics apart from education, self-employed status and risk attitude have little effect on the share of wealth invested in stocks.

Table 3: Coefficients and Marginal Effects for Participation, Normal Model for Transformed Share

	Mean Cf	Std.Cf	Prob(Cf>0)	Mean ME	Std. ME	Prob(ME>0)
const	-1.6969	0.2787				
age	0.0104	0.0106	0.84	0.0017	0.0007	1
age2	-0.0045	0.0106	0.33			
nw	0.0287	0.0021	1	0.0064	0.0009	1
nw2	-0.0121	0.0014	0			
nw3	0.0015	0.0002	1			
income	0.0004	0.012	0.51	-0.0002	0.0027	0.47
income2	-0.0143	0.0441	0.38			
eduB	0.3727	0.0656	1	0.1043	0.0219	1
eduD	0.2177	0.0541	1	0.057	0.0158	1
eduHS	0.3361	0.0845	1	0.0929	0.0269	1
olf	-0.0795	0.0842	0.17	-0.0218	0.0234	0.17
und	-0.2724	0.1697	0.05	-0.0665	0.0399	0.05
youngchild	0.0266	0.0253	0.85	0.0075	0.0072	0.85
selfempl	-0.0514	0.0794	0.26	-0.0142	0.0224	0.26
health	-0.1523	0.0619	0.01	-0.0407	0.0168	0.01
horizon 1	0.1865	0.0657	1	0.0577	0.0218	1
horizon 2	0.2202	0.054	1	0.0687	0.0186	1
risk	0.5338	0.0776	1	0.1817	0.0321	1
bizeq	-0.005	0.001	0	-0.0014	0.0003	0
housingeq	-0.0062	0.0015	0	-0.0018	0.0005	0
super w	0.0009	0.0024	0.65	0.0003	0.0007	0.65
super ret	0.002	0.0032	0.72	0.0006	0.0009	0.72
nesb	-0.3931	0.0671	0	-0.1287	0.0225	0
st nsw	0.3078	0.1266	0.99	0.0974	0.0376	0.99
st vic	0.2647	0.1261	0.98	0.0824	0.0373	0.98
st qld	0.2211	0.1282	0.96	0.0677	0.0379	0.96
st sa	0.3142	0.1393	0.99	0.0999	0.0428	0.99
st wa	0.2289	0.1363	0.96	0.0704	0.0409	0.96
st act	0.3625	0.204	0.96	0.1189	0.0687	0.96

APPENDIX

A1. The MCMC Algorithms for the Two Part Models with Normal and Mixture of Normals Disturbances

Two-Part Model with Normal Disturbances

The joint posterior distribution of parameters and latent data on the two part model with normal disturbances is proportional to the product of (6) and (7). To approximate this posterior distribution we construct the Gibbs sampling algorithm which iterates between the following four blocks:

1. Sample $I_i^*|\mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}$ for $i = 1, ..., n$. When $I_i = 1$, draw $I_i^*|(\mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}) \sim N(\boldsymbol{\beta}'\mathbf{z}_i, 1)$ truncated to $I_i^* > 0$;

   When $I_i = 0$, draw $I_i^*|(\mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}) \sim N(\boldsymbol{\beta}'\mathbf{z}_i, 1)$ truncated to $I_i^* < 0$.

2. $\boldsymbol{\beta}|(\mathbf{I}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}_{-\beta}) \sim N(\overline{\boldsymbol{\beta}}, \overline{H}_\beta)$ where $\overline{H}_\beta = \underline{H}_\beta + \mathbf{Z}'\mathbf{Z}$ and $\overline{\boldsymbol{\beta}} = \overline{H}_\beta^{-1}(\underline{H}_\beta\underline{\boldsymbol{\beta}} + \mathbf{Z}'\mathbf{I}^*)$.

3. $\boldsymbol{\alpha}|(\mathbf{I}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}_{-\alpha}) \sim N(\overline{\boldsymbol{\alpha}}, \overline{H}_\alpha)$ where $\overline{H}_\alpha = \underline{H}_\alpha + h\mathbf{X}'\mathbf{X}$ and $\overline{\boldsymbol{\alpha}} = \overline{H}_\alpha^{-1}(\underline{H}_\alpha\underline{\boldsymbol{\alpha}} + h\mathbf{X}'\mathbf{S})$.

4. $\overline{S}h|(\mathbf{I}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}_{-h}) \sim \chi^2(\overline{\nu})$ where $\overline{S} = \underline{S} + \sum_{i=1}^n (S_i - \boldsymbol{\alpha}'\mathbf{x}_i)^2$ and $\overline{\nu} = \underline{\nu} + n_{I=1}$.

Two Part Model with Mixture of Normals Disturbances

The joint posterior distribution of $\boldsymbol{\theta}$ and latent data in the two part model with mixture of normals disturbances is proportional to the product of (8) and (9). To approximate this posterior we construct a Gibbs sampling algorithm which iterates between the following six blocks:

1. Sample $I_i^*|\mathbf{S}, \mathbf{I}, \mathbf{Z}, \mathbf{s}, \boldsymbol{\theta}$ for $i = 1, ..., n$. When $I_i = 1$, draw $I_i^*|(\mathbf{S}, \mathbf{I}, \mathbf{Z}, \mathbf{s}, \boldsymbol{\theta}) \sim N(\boldsymbol{\beta}'\mathbf{z}_i, 1)$ truncated to $I_i^* > 0$;

   When $I_i = 0$, draw $I_i^*|(\mathbf{S}, \mathbf{I}, \mathbf{Z}, \mathbf{s}, \boldsymbol{\theta}) \sim N(\boldsymbol{\beta}'\mathbf{z}_i, 1)$ truncated to $I_i^* < 0$;

2. $P(s_i = j|\mathbf{S}, \mathbf{I}^*, \mathbf{I}, \mathbf{Z}, \mathbf{s}_{-i}, \boldsymbol{\theta}) \propto \pi_j\phi(S_i; \mu_j + \boldsymbol{\alpha}'\mathbf{x}_i, \sigma_{1j})$ where $\phi(a, B)$ denotes probability density function of normal distribution with mean a and variance B.

3. $\boldsymbol{\alpha}|(\mathbf{I}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \mathbf{s}, \boldsymbol{\theta}_{-\alpha}) \sim N(\overline{\boldsymbol{\alpha}}, \overline{H}_\alpha)$ where $\overline{H}_\alpha = \underline{H}_\alpha + \mathbf{Z}'\mathbf{Z}$ and $\overline{\boldsymbol{\alpha}} = \overline{H}_\alpha^{-1}(\underline{H}_\alpha\underline{\boldsymbol{\alpha}} + \mathbf{Z}'\mathbf{I}^*)$.

4. $\boldsymbol{\gamma}|(\mathbf{I}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \mathbf{s}, \boldsymbol{\theta}_{-\gamma}) \sim N(\overline{\boldsymbol{\gamma}}, \overline{H}_\gamma)$ where $\overline{H}_\gamma = \underline{H}_\gamma + \mathbf{W}'Q\mathbf{W}$ and $\overline{\boldsymbol{\gamma}} = \overline{H}_\gamma^{-1}(\underline{H}_\gamma\underline{\boldsymbol{\gamma}} + \mathbf{W}'Q\mathbf{S})$.

5. $\overline{S}_j h_j|(\mathbf{I}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \mathbf{s}, \boldsymbol{\theta}_{-h_j}) \sim \chi^2(\overline{\nu}_j)$ where $\overline{S}_j = \underline{S}_j + \sum_{i=1}^n \iota(s_i = j)(S_i^* - \mu_{s_i} - \boldsymbol{\alpha}'\mathbf{x}_i)^2$ and $\overline{\nu}_j = \underline{\nu}_j + n_{I=1}^j$.

6. $\boldsymbol{\pi}|(\mathbf{I}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}_{-\boldsymbol{\pi}}) \sim Dirichlet(r + n_{I=1}^1, ..., r + n_{I=1}^m)$.

A2. Marginal Effects

The expressions for the marginal effects (14)-(20) in the two part models with normal and mixture of normals disturbances can be derived as follows. In both two part models the conditional probability of stock market participation $Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A)$ is equal to the standard normal cdf evaluated at $\boldsymbol{\beta}'\mathbf{z}_i$:

$$Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A) = \Phi(\boldsymbol{\beta}'\mathbf{z}_i) \tag{A-1}$$

where $\Phi(a)$ denotes the standard normal cdf evaluated at a. Then the marginal effect of a continuous variable z_{ki} on probability of stock market participation of individual i is:

$$MEP_{z_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, N = \frac{\partial\Phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\partial z_{ki}} = \beta_k \cdot \phi(\boldsymbol{\beta}'\mathbf{z}_i), \tag{A-2}$$

where $\phi(a)$ denotes standard normal pdf evaluated at a. The marginal effect of a discrete z_{ki} on probability of stock market participation of individual i is:

$$\begin{aligned} MEP_{z_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, N &= \Phi(\beta_k(z_{ki}+1) + \boldsymbol{\beta}_{-\beta_k}\mathbf{z}_{-z_{ki}}) - \Phi(\boldsymbol{\beta}'\mathbf{z}_i) \\ &= \Phi(\beta_k + \boldsymbol{\beta}'\mathbf{z}_i) - \Phi(\boldsymbol{\beta}'\mathbf{z}_i). \end{aligned} \tag{A-3}$$

The expected values of the fraction of wealth invested in shares conditional on participation in normal and mixture of normals two part models $E(s_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, N)$ and $E(s_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, M)$ are defined as follows:

$$E(s_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, N) = \int_{-\infty}^{+\infty} \frac{exp(\boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i})}{1 + exp(\boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i})} \cdot \phi(\varepsilon_{1i}|0, \sigma_1) d\varepsilon_{i1}, \tag{A-4}$$

and

$$E(s_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, M) = \int_{-\infty}^{+\infty} \frac{exp(\boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i})}{1 + exp(\boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i})} \cdot \Sigma_{j=1}^m \pi_j \phi(\varepsilon_{1i}|\mu_j, \sigma_{1j}) d\varepsilon_{i1}. \tag{A-5}$$

where $\phi(.|a, B)$ denotes probability density function of normal distribution with mean a and variance B. Because the closed-form expressions for conditional expectations (A-4) and (A-5) do not exist, we approximate these integrals using Monte-Carlo integration. To approximate marginal effects of variable x_{ki} we compute the difference between these integrals evaluated at $x_{ki} + 1$ and x_{ki} for both discrete and continuous x_{ki}, because increase by one represents a fairly small change for continuous covariates in our data.

Finally, the unconditional expectations of the observed fraction of wealth invested in shares in normal and mixture of normals two part models $E(s_i^o|\mathbf{z}_i, \boldsymbol{\theta}, N)$ and $E(s_i^o|\mathbf{z}_i, \boldsymbol{\theta}, M)$ are defined as follows:

$$E(s_i^o|\mathbf{z}_i, \boldsymbol{\theta}, N) = \Phi(\boldsymbol{\beta}'\mathbf{z}_i) \cdot \int_{-\infty}^{+\infty} \frac{exp(\boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i})}{1 + exp(\boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i})} \cdot \phi(\varepsilon_{1i}|0, \sigma_1) d\varepsilon_{i1}, \tag{A-6}$$

and

$$E(s_i^o|\mathbf{z}_i,\boldsymbol{\theta},M) = \Phi(\boldsymbol{\beta}'\mathbf{z}_i)\int_{-\infty}^{+\infty}\frac{exp(\boldsymbol{\alpha}'\mathbf{x}_i+\varepsilon_{1i})}{1+exp(\boldsymbol{\alpha}'\mathbf{x}_i+\varepsilon_{1i})}\cdot(\Sigma_{j=1}^m\pi_j\phi(\varepsilon_{1i}|\mu_j,\sigma_{1j}))d\varepsilon_{i1}. \tag{A-7}$$

The closed-form expressions for conditional expectations (A-6) and (A-7) also do not exist. We approximate these integrals using Monte-Carlo integration and approximate marginal effects of variable x_{ki} by computing the difference between these integrals evaluated at $x_{ki}+1$ and x_{ki} for both discrete and continuous x_{ki}.

In the tobit model the expressions for the marginal effects are defined as follows. The conditional probability of stock market participation $Prob(I_i^*>0|\mathbf{z}_i,\boldsymbol{\theta},T)$ is equal to the standard normal cdf evaluated at $\boldsymbol{\beta}'\mathbf{z}_i$:

$$Prob(I_i^*>0|\mathbf{z}_i,\boldsymbol{\theta},T) = \Phi(\boldsymbol{\beta}'\mathbf{z}_i) \tag{A-8}$$

and marginal effects of continuous and discreet variables on this probability are as defined in (A-2) and (A-3). The expected values of the fraction of wealth invested in shares conditional on participation in the tobit model $E(s_i^*|s_i^*>0,\mathbf{z}_i,\boldsymbol{\theta},T)$ is defined as follows:

$$E(s_i^*|s_i^*>0,\mathbf{z}_i,\boldsymbol{\theta},T) = \boldsymbol{\beta}'\mathbf{z}_i+\sqrt{\sigma}\frac{\phi_{SN}(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})}{\Phi(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})}. \tag{A-9}$$

The marginal effect of a continuous variable z_{ki} on this expectation is given by:

$$\begin{aligned} MES_{z_{ki}^c}|\mathbf{z}_i,\boldsymbol{\theta},T &= \frac{\partial E(s_i^*|s_i^*>0,\mathbf{z}_i,\boldsymbol{\theta},T)}{\partial z_{ki}} \\ &= \beta_k(1-\frac{(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})\phi_{SN}(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})}{\Phi(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})}-\left(\frac{\phi_{SN}(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})}{\Phi(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})}\right)^2). \end{aligned} \tag{A-10}$$

The marginal effect of a discrete variable z_{ki} on this expectation is given by the difference between expression (A-11) evaluated at $z_{ki}+1$ and z_{ki}.

The expected values of the fraction of wealth invested in shares in the tobit model $E(s_i^o|\mathbf{z}_i,\boldsymbol{\theta},T)$ is derived in Wooldridge (2002) Ch.16, and is defined as follows:

$$\begin{aligned} E(s_i^o|\mathbf{z}_i,\boldsymbol{\theta},T) &= P(s_i^*>0|\mathbf{z}_i,\boldsymbol{\theta},T)\cdot E(s_i|s_i^*>0,\mathbf{z}_i,\boldsymbol{\theta},T) \\ &= \Phi(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})\boldsymbol{\beta}'\mathbf{z}_i+\sqrt{\sigma}\phi(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}}). \end{aligned} \tag{A-11}$$

The marginal effect of a continuous variable z_{ki} on this expectation is given by:

$$MESU_{z_{ki}^c}|\mathbf{z}_i,\boldsymbol{\theta},T = \frac{\partial E(s_i^o|\mathbf{z}_i,\boldsymbol{\theta},T)}{\partial z_{ki}} = \Phi(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma}})\beta_k. \tag{A-12}$$

In the normal two-part model for untransformed share marginal effects of covariates on participation probability are defined as in (A-2) and (A-3), marginal effects of covariates on share conditional on participation are just coefficients in share equation α_k and marginal effects of covariates on share (unconditional of participation) are given by $\phi(\boldsymbol{\beta}'\mathbf{z}_i) \cdot \boldsymbol{\alpha}'\mathbf{x}_i \cdot \beta_k + \Phi(\boldsymbol{\beta}'\mathbf{z}_i)\alpha_k$.

A3. Specification of Prior Distributions

We specify the following hyper-parameters of the prior distribution of $\boldsymbol{\theta}$ in the normal model for transformed share:

1. The mean of the prior distribution of the vector of coefficients $[\underline{\boldsymbol{\alpha}}', \underline{\boldsymbol{\beta}}']'$ and the precision of this distribution $\underline{H}$ are specified as follows:

$$\begin{aligned} \underline{\boldsymbol{\alpha}} &= [.38, \mathbf{0}'_{K_x \times 1}]', \\ \underline{\boldsymbol{\beta}} &= [-.23, \mathbf{0}'_{K_z \times 1}]', \end{aligned}$$

$$\underline{H} = \begin{bmatrix} \underline{H}_\alpha & \mathbf{0} \\ \mathbf{0} & \underline{H}_\beta \end{bmatrix},$$

   $\underline{H}_\alpha = (1/50)\mathbf{I}_{K_x \times K_x}$ and $\underline{H}_\beta = (1/50)\mathbf{I}_{K_z \times K_z}$. The priors of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are diffuse and are specified so that the prior distributions $p(S_i^*|\mathbf{x}_i)$ and $p(I_i|\mathbf{z}_i)$ for $i = 1, ..., n$ are centered at sample means of $S_i|I_i = 1$ and I_i respectively.

2. The hyper-parameters which govern the prior distribution of the parameters of the distribution of ε_{1i}, $\underline{S}$ and $\underline{\nu}$ are specified as follows:

$$\underline{S} = 10, \ \underline{\nu} = 3.$$

   This prior distribution centers variance of $S_i^*|\mathbf{x}_i, \boldsymbol{\theta}$ around sample variance of $S_i|I_i = 1$.

In the mixture of normals two-part model for transformed share the following hyper-parameters of the prior distribution of $\boldsymbol{\theta}$ are used:

1. The mean of the prior distribution of the vector of coefficients $[\underline{\boldsymbol{\mu}}', \underline{\boldsymbol{\alpha}}', \underline{\boldsymbol{\beta}}']'$ and the precision of this distribution $\underline{H}_m$ are specified as follows:

$$\begin{aligned} \underline{\mu} &= \mathbf{0}_{m \times 1}, \\ \underline{\boldsymbol{\alpha}} &= [0.38, \mathbf{0}'_{K_x \times 1}]', \\ \underline{\boldsymbol{\beta}} &= [-0.23, \mathbf{0}'_{K_z \times 1}]', \end{aligned}$$

$$\underline{H}_m = \begin{bmatrix} \underline{H}_\mu & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \underline{H}_\alpha & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \underline{H}_\beta \end{bmatrix},$$

where $\underline{H}_\mu = .1\mathbf{I}_{m\times M}$, $\underline{H}_\alpha = (1/50)\mathbf{I}_{K_x\times K_x}$, $\underline{H}_\beta = (1/50)\mathbf{I}_{K_z\times K_z}$. In this prior, low precision of $\boldsymbol{\mu}_1$ implies substantial probability of multimodality in the conditional on parameters and $\mathbf{x}_i$ distribution of S_i^*. The priors of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are diffuse and are specified so that the prior distributions $p(S_i^*|\mathbf{x}_i)$ and $p(I_i|\mathbf{z}_i)$ for $i = 1, ..., n$ are centered at sample means of $S_i|I_i = 1$ and I_i respectively.

2. The hyper-parameters which govern the prior distribution of the parameters of the distribution of $\varepsilon)1i$, $\underline{S}_j$ and $\underline{\nu}_j$, $j = 1, ..., m$, are specified as follows:

$$\underline{S}_j = 10, \ \underline{\nu}_j = 3.$$

This prior distribution centers variance of $S_i^*|\mathbf{x}_i, \boldsymbol{\theta}, s_i = j$ around sample variance of $S_i|I_i = 1$ for $j = 1, ..., m$.

3. The parameters of the prior distribution of the marginal probabilities of mixture components $\boldsymbol{\pi}$, $r_1, ..., r_m$ are all set to 1.

In the tobit model the following hyper-parameters of the prior distribution of $\boldsymbol{\theta}$ are used:

1. The mean of the prior distribution of the vector of coefficients $\underline{\boldsymbol{\beta}}'$ and the precision of this distribution $\underline{H}_\beta$ are specified as follows:

$$\underline{\boldsymbol{\beta}} = [0.26, \mathbf{0}'_{K_z\times 1}]',$$

and $\underline{H}_\beta = (1/50)\mathbf{I}_{K_z\times K_z}$. The prior $\boldsymbol{\beta}$ is diffuse and are specified so that the prior distributions $p(s_i^*|\mathbf{x}_i)$ for $i = 1, ..., n$ are centered at sample means of s_i^o.

2. The hyper-parameters which govern the prior distribution of the parameters of the distribution of ε_{1i}, $\underline{S}_j$ and $\underline{\nu}_j$, $j = 1, ..., m$, are specified as follows:

$$\underline{S}_j = 2, \ \underline{\nu}_j = 3.$$

In the normal two-part model for transformed share the following hyper-parameters of the prior distribution are specified:

1. The mean of the prior distribution of the vector of coefficients $[\underline{\boldsymbol{\alpha}}', \underline{\boldsymbol{\beta}}']'$ and the precision of this distribution $\underline{H}$ are specified as follows:

$$\underline{\boldsymbol{\alpha}} = [.55, \mathbf{0}'_{K_x\times 1}]',$$

$$\underline{\boldsymbol{\beta}} = [-.23, \mathbf{0}'_{K_z \times 1}]',$$

$$\underline{H} = \begin{bmatrix} \underline{H}_\alpha & \mathbf{0} \\ \mathbf{0} & \underline{H}_\beta \end{bmatrix},$$

$\underline{H}_\alpha = (1/50)\mathbf{I}_{K_x \times K_x}$ and $\underline{H}_\beta = (1/50)\mathbf{I}_{K_z \times K_z}$. The priors of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are diffuse and are specified so that the prior distributions $p(s_i^*|\mathbf{x}_i)$ and $p(I_i|\mathbf{z}_i)$ for $i = 1, ..., n$ are centered at sample means of $s_i^o|I_i = 1$ and I_i respectively.

2. The hyper-parameters which govern the prior distribution of the parameters of the distribution of ε_{1i}, $\underline{S}$ and $\underline{\nu}$ are specified as follows:

$$\underline{S} = 2,\ \underline{\nu} = 3.$$

References

C. Bertaut. Stockholding behavior of U.S. households: evidence from the 1983-1989 Survey of Consumer Finances. *Review of Economics and Statistics*, 80(2):263–275, 1998.

J. Campbell. Household protfolios. *Journal of Finance*, 61(4):1553–1604, 2006.

S. Chib. Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association*, 90:1313–1321, 1995.

J. Cocco, Gomes F., and Maenhout P. Consumption and portfolio choice over the life cycle. *Review of Financial Studies*, 18:491533, 2005.

J. Geweke. Getting it right: Joint distribution tests of posterior simulators. *Journal of the American Statistical Association*, 99:799–804, 2004.

J. Geweke. Interpretation and inference in mixture models: simple MCMC works. *Computational Statistics and Data Analysis*, 51:3529–2550, 2007.

M. Giannetti and Y. Koskinen. Investor protection, equity returns, and financial globalization. *Working Paper*, 2007.

L. Guiso, T. Jappelli, and D. Terlizzese. Income risk, borrowing constraints and portfolio choice. *American Economic Review*, 86:158172, 1996.

L. Guiso, M. Haliassos, and T. Jappelli. *Household Portfolios.* MIT Press, Cambridge, MA, 2002.

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

S. Hochguertel. Precautionary motives and portfolio decisions. *Journal of Applied Econometrics*, 18:61–77, 2003.

R. Kieschnick and BD. McCullough. Regression analysis of variates observed on (0, 1): percentages, proportions and fractions. *Statistical Modelling*, 3:193213, 2003.

G. Koop, DJ. Poirier, and JL. Tobias. *Bayesian Econometric Methods.* Cambridge University Press, 2007.

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

M. Paliela. Foregone gains of incomplete portfolios. *Review of Financial Studies*, forthcoming, 2007.

J. Tobin. Liquidity preference as behavior towards risk. *Review of Economic Studies*, 25:65–86, 1958.

A. Vissing-Jorgensen. *Perspectives on Behavioral Finance: does Irrationality Disappear with Wealth? Evidence from Expectations and Actions.* NBER Macroeconomics Annual 2003. MIT Press, 2004.

M. Wooden and N. Watson. The HILDA survey and its contribution to economic and social research (so far). *The Economic Record*, 83:208–231, 2007.

J. Wooldridge. *Econometric Analysis of Cross Section and Panel Data.* MIT Press, 2002.

Table 4: Coefficients and Marginal Effects for Participation, Tobit Model

	Mean Cf	**Std.Cf**	**Prob(Cf>0)**	**Mean ME**	**Std. ME**	**Prob(ME>0)**
const	-0.7769	0.1422	0			
age	0.0068	0.0053	0.9	0.0013	0.0007	0.98
age2	-0.0047	0.0052	0.18			
nw	0.0127	0.0009	1	0.0055	0.0006	1
nw2	-0.0059	0.0006	0			
nw3	0.0008	0.0001	1			
income	-0.0057	0.0054	0.15	-0.0029	0.0025	0.12
income2	0.0091	0.0184	0.69			
eduB	0.1911	0.0306	1	0.1086	0.0195	1
eduD	0.1074	0.0265	1	0.0581	0.015	1
eduHS	0.1623	0.0397	1	0.091	0.0243	1
olf	-0.0058	0.0391	0.44	-0.0031	0.0226	0.44
und	-0.0941	0.0837	0.13	-0.0491	0.0436	0.13
youngchild	0.019	0.0119	0.95	0.0108	0.0068	0.95
selfempl	-0.0707	0.0369	0.03	-0.039	0.0204	0.03
health	-0.0815	0.0296	0	-0.0447	0.0163	0
horizon 1	0.0553	0.0302	0.96	0.033	0.0185	0.96
horizon 2	0.0932	0.0246	1	0.0564	0.0155	1
risk	0.259	0.0317	1	0.1668	0.0233	1
bizeq	-0.0019	0.0004	0	-0.0011	0.0002	0
housingeq	-0.0021	0.0006	0	-0.0012	0.0004	0
super w	-0.0005	0.0009	0.26	-0.0003	0.0005	0.26
super ret	0.0006	0.001	0.7	0.0003	0.0006	0.7
nesb	-0.1874	0.0333	0	-0.1177	0.0205	0
st nsw	0.1521	0.0658	0.99	0.0936	0.0386	0.99
st vic	0.1631	0.0659	0.99	0.1009	0.0387	0.99
st qld	0.1509	0.0669	0.99	0.0929	0.0395	0.99
st sa	0.145	0.0706	0.98	0.0891	0.0422	0.98
st wa	0.1388	0.0703	0.98	0.085	0.0418	0.98
st act	0.1543	0.0958	0.95	0.0961	0.0599	0.95

Table 5: Coefficients and **Conditional** Marginal Effects, Share of Risky Assets, Normal Model for Transformed Share

	Mean Cf	**Std.Cf**	**Prob(Cf>0)**	**Mean ME**	**Std. ME**	**Prob(ME>0)**
const	1.028	0.5429	0.97			
age	-0.0155	0.0225	0.24	-0.001	0.0019	0.28
age2	0.0094	0.0224	0.67			
nw	-0.0013	0.004	0.37	-0.0001	0.0019	0.47
nw2	0.0008	0.0025	0.63			
nw3	-0.0001	0.0004	0.41			
income	-0.0796	0.0229	0	-0.0093	0.0032	0
income2	0.2002	0.0755	1			
eduB	0.3364	0.135	0.99	0.0509	0.0207	0.99
eduD	0.1512	0.1238	0.89	0.0231	0.0193	0.88
eduHS	0.2199	0.1802	0.89	0.0333	0.0274	0.89
olf	0.1981	0.1725	0.88	0.0299	0.0247	0.89
und	0.59	0.3925	0.94	0.0881	0.0543	0.95
youngchild	0.0443	0.0537	0.8	0.0064	0.0082	0.79
selfempl	-0.5601	0.1602	0	-0.0833	0.0245	0
health	-0.1146	0.1385	0.21	-0.0179	0.0207	0.2
horizon 1	-0.26	0.1343	0.03	-0.0399	0.0208	0.03
horizon 2	-0.0133	0.107	0.45	-0.002	0.0167	0.46
risk	0.6407	0.1296	1	0.0927	0.0186	1
bizeq	-0.0037	0.0016	0.01	-0.0005	0.0018	0.39
housingeq	-0.0017	0.0024	0.24	-0.0003	0.0019	0.43
super w	0.0024	0.0035	0.76	0.0005	0.0019	0.6
super ret	0.0041	0.0041	0.84	0.0005	0.0019	0.62

Table 6: Coefficients and **Conditional** Marginal Effects, Share of Risky Assets, Tobit Model

	Mean Cf	Std.Cf	Prob(Cf>0)	Mean ME	Std. ME	Prob(ME>0)
const	-0.7769	0.1422	0			
age	0.0068	0.0053	0.9	0.0006	0.0003	0.98
age2	-0.0047	0.0052	0.18			
nw	0.0127	0.0009	1	0.0024	0.0003	1
nw2	-0.0059	0.0006	0			
nw3	0.0008	0.0001	1			
income	-0.0057	0.0054	0.15	-0.0012	0.0011	0.12
income2	0.0091	0.0184	0.69			
eduB	0.1911	0.0306	1	0.0472	0.0084	1
eduD	0.1074	0.0265	1	0.0253	0.0065	1
eduHS	0.1623	0.0397	1	0.0395	0.0105	1
olf	-0.0058	0.0391	0.44	-0.0013	0.0098	0.44
und	-0.0941	0.0837	0.13	-0.0216	0.0192	0.13
youngchild	0.019	0.0119	0.95	0.0048	0.0031	0.95
selfempl	-0.0707	0.0369	0.03	-0.0169	0.0088	0.03
health	-0.0815	0.0296	0	-0.0194	0.007	0
horizon 1	0.0553	0.0302	0.96	0.0143	0.008	0.96
horizon 2	0.0932	0.0246	1	0.0245	0.0068	1
risk	0.259	0.0317	1	0.0746	0.0119	1
bizeq	-0.0019	0.0004	0	-0.0005	0.0001	0
housingeq	-0.0021	0.0006	0	-0.0005	0.0002	0
super w	-0.0005	0.0009	0.26	-0.0001	0.0002	0.26
super ret	0.0006	0.001	0.7	0.0001	0.0003	0.7
nesb	-0.1874	0.0333	0	-0.0516	0.0092	0
st nsw	0.1521	0.0658	0.99	0.0406	0.0167	0.99
st vic	0.1631	0.0659	0.99	0.0439	0.0168	0.99
st qld	0.1509	0.0669	0.99	0.0403	0.0171	0.99
st sa	0.145	0.0706	0.98	0.0387	0.0183	0.98
st wa	0.1388	0.0703	0.98	0.0368	0.0182	0.98
st act	0.1543	0.0958	0.95	0.042	0.0265	0.95

Table 7: Coefficients and **Unconditional** Marginal Effects, Share of Risky Assets, Normal Model for Transformed Share

	Mean Cf	**Std.Cf**	**Prob(Cf>0)**	**Mean ME**	**Std. ME**	**Prob(ME>0)**
const	1.028	0.5429	0.97			
age	-0.0155	0.0225	0.24	0.0007	0.0016	0.68
age2	0.0094	0.0224	0.67			
nw	-0.0013	0.004	0.37	0.0036	0.0017	0.99
nw2	0.0008	0.0025	0.63			
nw3	-0.0001	0.0004	0.41			
income	-0.0796	0.0229	0	-0.0021	0.0022	0.17
income2	0.2002	0.0755	1			
eduB	0.3364	0.135	0.99	0.0688	0.0138	1
eduD	0.1512	0.1238	0.89	0.0354	0.0101	1
eduHS	0.2199	0.1802	0.89	0.0578	0.0163	1
olf	0.1981	0.1725	0.88	-0.0064	0.0144	0.31
und	0.59	0.3925	0.94	-0.0243	0.0265	0.17
youngchild	0.0443	0.0537	0.8	0.0057	0.0052	0.87
selfempl	-0.5601	0.1602	0	-0.0243	0.0134	0.03
health	-0.1146	0.1385	0.21	-0.0258	0.01	0
horizon 1	-0.26	0.1343	0.03	0.0213	0.0122	0.97
horizon 2	-0.0133	0.107	0.45	0.0377	0.0111	1
risk	0.6407	0.1296	1	0.1373	0.0242	1
risk nr	-0.129	0.4335	0.38	0	0	0
bizeq	-0.0037	0.0016	0.01	-0.0009	0.0015	0.27
housingeq	-0.0017	0.0024	0.24	-0.001	0.0015	0.27
super w	0.0024	0.0035	0.76	0.0002	0.0016	0.55
super ret	0.0041	0.0041	0.84	0.0005	0.0017	0.59

Table 8: Coefficients and **Unconditional** Marginal Effects, Share of Risky Assets, Tobit Model

	Mean Cf	**Std.Cf**	**Prob(Cf>0)**	**Mean ME**	**Std. ME**	**Prob(ME>0)**
const	-0.7769	0.1422	0			
age	0.0068	0.0053	0.9	0.0006	0.0003	0.98
age2	-0.0047	0.0052	0.18			
nw	0.0127	0.0009	1	0.0025	0.0005	1
nw2	-0.0059	0.0006	0			
nw3	0.0008	0.0001	1			
income	-0.0057	0.0054	0.15	-0.0013	0.0012	0.12
income2	0.0091	0.0184	0.69			
eduB	0.1911	0.0306	1	0.0499	0.0114	1
eduD	0.1074	0.0265	1	0.0254	0.0075	1
eduHS	0.1623	0.0397	1	0.0412	0.0128	1
olf	-0.0058	0.0391	0.44	-0.0013	0.0105	0.44
und	-0.0941	0.0837	0.13	-0.021	0.0191	0.13
youngchild	0.019	0.0119	0.95	0.0052	0.0035	0.95
selfempl	-0.0707	0.0369	0.03	-0.0173	0.0093	0.03
health	-0.0815	0.0296	0	-0.0197	0.0076	0
horizon 1	0.0553	0.0302	0.96	0.0158	0.0093	0.96
horizon 2	0.0932	0.0246	1	0.0275	0.0086	1
risk	0.259	0.0317	1	0.0904	0.0183	1
bizeq	-0.0019	0.0004	0	-0.0005	0.0001	0
housingeq	-0.0021	0.0006	0	-0.0006	0.0002	0
super w	-0.0005	0.0009	0.26	-0.0001	0.0002	0.26
super ret	0.0006	0.001	0.7	0.0001	0.0003	0.7
nesb	-0.1874	0.0333	0	-0.0605	0.0122	0
st nsw	0.1521	0.0658	0.99	0.0464	0.0186	0.99
st vic	0.1631	0.0659	0.99	0.0504	0.0187	0.99
st qld	0.1509	0.0669	0.99	0.046	0.0191	0.99
st sa	0.145	0.0706	0.98	0.0441	0.0207	0.98
st wa	0.1388	0.0703	0.98	0.0418	0.0204	0.98
st act	0.1543	0.0958	0.95	0.0488	0.0315	0.95

Table 9: Coefficients and **Conditional** Marginal Effects, Share of Risky Assets, Mixture Model for Transformed Share

	Mean Cf	**Std.Cf**	**Prob(Cf>0)**	**Mean ME**	**Std. ME**	**Prob(ME>0)**
const 1	-0.7664	2.3357	0.38			
const 2	0.9523	2.4197	0.64			
const	1.6643	2.3888	0.75			
age	-0.0164	0.0257	0.26	-0.0008	0.0023	0.36
age2	0.0111	0.025	0.67			
nw	-0.0018	0.0041	0.33	-0.0001	0.0023	0.48
nw2	0.0011	0.0025	0.66			
nw3	-0.0001	0.0004	0.38			
income	-0.074	0.023	0	-0.0088	0.0033	0
income2	0.1869	0.0746	0.99			
eduB	0.3325	0.1358	0.99	0.0516	0.0209	0.99
eduD	0.1508	0.1229	0.89	0.0238	0.0192	0.88
eduHS	0.2159	0.1788	0.89	0.0343	0.0268	0.9
olf	0.2329	0.1776	0.91	0.0355	0.0273	0.91
und	0.6194	0.4463	0.92	0.0872	0.0652	0.91
youngchild	0.0503	0.0525	0.83	0.0075	0.0086	0.82
selfempl	-0.502	0.1617	0	-0.0773	0.0249	0
health	-0.1291	0.1395	0.18	-0.0189	0.0219	0.2
horizon 1	-0.2505	0.1313	0.03	-0.0375	0.0199	0.03
horizon 2	0.0005	0.1053	0.5	0.0011	0.0164	0.54
risk	0.6526	0.128	1	0.0977	0.0189	1
bizeq	-0.0039	0.0017	0.01	-0.0006	0.0022	0.38
housingeq	-0.0019	0.0025	0.22	-0.0003	0.0022	0.47
super w	0.002	0.0035	0.72	0.0003	0.0022	0.56
super ret	0.0042	0.0041	0.85	0.0007	0.0022	0.61

Table 10: Coefficients and **Conditional** Marginal Effects, Share of Risky Assets, Normal Model for Untransformed Share

	Mean Cf	Std.Cf	Prob(Cf>0)	Mean ME	Std. ME	Prob(ME>0)
const	0.6305	0.0985	1			
age	-0.0018	0.004	0.32	-0.0009	0.0008	0.13
age2	0.001	0.0039	0.6			
nw	-0.0002	0.0006	0.39	-0.0001	0.0004	0.42
nw2	0.0002	0.0004	0.66			
nw3	0	0.0001	0.39			
income	-0.0111	0.0035	0	-0.0089	0.0028	0
income2	0.0286	0.0116	0.99			
eduB	0.0501	0.0209	0.99	0.0501	0.0209	0.99
eduD	0.0203	0.019	0.86	0.0203	0.019	0.86
eduHS	0.0291	0.0279	0.85	0.0291	0.0279	0.85
olf	0.0481	0.0269	0.96	0.0481	0.027	0.96
und	0.054	0.0659	0.8	0.054	0.0659	0.8
youngchild	0.011	0.0082	0.91	0.011	0.0082	0.91
selfempl	-0.0718	0.0248	0	-0.0718	0.0248	0
health	-0.0241	0.0216	0.13	-0.0241	0.0216	0.13
horizon 1	-0.0367	0.0206	0.04	-0.0367	0.0206	0.04
horizon 2	0.001	0.0162	0.52	0.001	0.0162	0.52
risk	0.1059	0.0201	1	0.1059	0.0201	1
bizeq	-0.0008	0.0002	0	-0.0008	0.0002	0
housingeq	-0.0004	0.0004	0.12	-0.0004	0.0004	0.12
super w	0.0001	0.0005	0.61	0.0001	0.0005	0.61
super ret	0.0004	0.0006	0.73	0.0004	0.0006	0.73

Table 11: Coefficients and **Unconditional** Marginal Effects, Share of Risky Assets, Mixture Model for Transformed Share

	Mean Cf	Std.Cf	Prob(Cf>0)	Mean ME	Std. ME	Prob(ME>0)
const 1	-0.7664	2.3357	0.38			
const 2	0.9523	2.4197	0.64			
const	1.6643	2.3888	0.75			
age	-0.0164	0.0257	0.26	0.0008	0.0021	0.63
age2	0.0111	0.025	0.67			
nw	-0.0018	0.0041	0.33	0.0033	0.0022	0.93
nw2	0.0011	0.0025	0.66			
nw3	-0.0001	0.0004	0.38			
income	-0.074	0.023	0	-0.0017	0.0026	0.26
income2	0.1869	0.0746	0.99			
eduB	0.3325	0.1358	0.99	0.0658	0.0139	1
eduD	0.1508	0.1229	0.89	0.0331	0.0096	1
eduHS	0.2159	0.1788	0.89	0.0565	0.0165	1
olf	0.2329	0.1776	0.91	-0.0046	0.0141	0.37
und	0.6194	0.4463	0.92	-0.0239	0.0257	0.16
youngchild	0.0503	0.0525	0.83	0.0058	0.0048	0.89
selfempl	-0.502	0.1617	0	-0.0216	0.0123	0.04
health	-0.1291	0.1395	0.18	-0.0239	0.0096	0.01
horizon 1	-0.2505	0.1313	0.03	0.0211	0.0121	0.97
horizon 2	0.0005	0.1053	0.5	0.0369	0.0108	1
risk	0.6526	0.128	1	0.136	0.0246	1
bizeq	-0.0039	0.0017	0.01	-0.001	0.0021	0.32
housingeq	-0.0019	0.0025	0.22	-0.001	0.0021	0.32
super w	0.002	0.0035	0.72	0.0002	0.0021	0.56
super ret	0.0042	0.0041	0.85	0.0004	0.0022	0.57

Table 12: Coefficients and **Unconditional** Marginal Effects, Share of Risky Assets, Normal Model for Untransformed Share

	Mean Cf	**Std.Cf**	**Prob(Cf>0)**	**Mean ME**	**Std. ME**	**Prob(ME>0)**
const	0.6305	0.0985	1	0	0	0
age	-0.0018	0.004	0.32	0.0008	0.0004	0.97
age2	0.001	0.0039	0.6			
nw	-0.0002	0.0006	0.39	0.0035	0.0005	1
nw2	0.0002	0.0004	0.66			
nw3	0	0.0001	0.39			
income	-0.0111	0.0035	0	-0.002	0.0017	0.11
income2	0.0286	0.0116	0.99			
eduB	0.0501	0.0209	0.99	0.0683	0.0144	1
eduD	0.0203	0.019	0.86	0.0343	0.0097	1
eduHS	0.0291	0.0279	0.85	0.0563	0.0166	1
olf	0.0481	0.0269	0.96	-0.003	0.0147	0.41
und	0.054	0.0659	0.8	-0.0285	0.0257	0.13
youngchild	0.011	0.0082	0.91	0.0066	0.0047	0.93
selfempl	-0.0718	0.0248	0	-0.0222	0.0129	0.04
health	-0.0241	0.0216	0.13	-0.0267	0.0102	0
horizon 1	-0.0367	0.0206	0.04	0.0216	0.0126	0.96
horizon 2	0.001	0.0162	0.52	0.0382	0.0112	1
risk	0.1059	0.0201	1	0.1415	0.0243	1
bizeq	-0.0008	0.0002	0	-0.001	0.0002	0
housingeq	-0.0004	0.0004	0.12	-0.0011	0.0003	0
super w	0.0001	0.0005	0.61	0.0002	0.0004	0.67
super ret	0.0004	0.0006	0.73	0.0004	0.0005	0.78