Background Risk and Household Portfolio Choice: Bayesian Analysis of a Generalized Selection Model

Olena Stavrunova* Oleg Yerokhin
UTS UTS

This Draft: May 2008

Abstract

This paper studies determinants of household portfolio decisions using recent data from Australia. A flexible econometric approach based on the finite mixture of normals sample selection model is used to analyze jointly the decision to participate in the market for risky assets and how much to invest in such assets. To study factors influencing these decisions the paper derives expressions for marginal effects of covariates in the mixture of normals sample selection model and evaluates their posterior distributions. We find that in our sample the mixture model outperforms normal model according to the marginal likelihood criterion, and that up to five mixture components are required to approximate the joint distribution of errors in the sample selection model. Results based on the preferred mixture model show that household characteristics and background risk variables have significant effect on participation decision, while having moderate predictive power with respect to the share of wealth invested in risky assets. The paper also demonstrates that relatively modest participation cost can rationalize prevalence of non-participation in the market for risky assets, with the cost in the range of 500-700 dollars explaining 90% of non-participation in our sample.

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 effects of fixed cost, background risk and other characteristics on stock market participation and household exposure to risky assets (public equity) 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 background risk, planning horizon, risk attitudes and other characteristics and study their impact on the structure of household portfolios. We find that controlling for wealth level and demographic characteristics, households in which head has higher subjective risk of a job loss 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 common stock. These findings confirm the results obtained by Guiso, Jappelli and Terlizzese (1996) and Hochguertel (2003) who studied the effect of subjective expectations about such risk on portfolio choice. Similar to other studies we also document strong impact of education, age, risk attitudes and net worth levels on decision to hold public equity.

The methodological novelty of the paper consists in addressing important properties of the household portfolio data in a systematic fashion. Guiso et al (2002) point out that the sample selection model provides a convenient reduced form framework to study household portfolio choices: it allows treating participation and allocation as two distinct decisions while simultaneously controlling for the unobserved factors which might affect both choices. However, it is well know that inference in the normal sample selection model is unreliable when the assumption of normality is violated. In our data the conditional distribution of the share of wealth invested in stocks among participating households is bimodal and platykurtic, which suggests that the normality assumption may not be satisfied. Hence, we employ the discrete mixture of normals selection model, which is flexible enough to accommodate any deviation from normality in the data (Ferguson, 1983). The paper takes Bayesian approach to inference and employs Markov Chain Monte Carlo simulation algorithm to access the joint posterior distribution of the parameters of the model. The method proposed by Chib (1995) is used to evaluate model's marginal likelihood and select the number of components in the mixture distribution. This selection procedure favors the five-component mixture of normals model.

The paper also uses data on individual financial wealth to obtain the posterior distribution of the magnitude of forgone earnings from non-participation in the stock market using method similar to that proposed by Vissing-Jorgensen (2004). With some additional assumptions we use this posterior distribution to construct the distribution of the fixed cost of stock market participation for non-participating households. The posterior distribution of the participation cost obtained from the five-component mixture model differs substantially from the posterior distribution generated by the normal model. In particular, normal model systematically predicts lower participation costs than five-component mixture model. These results show the importance of accommodating non-normality in the data for drawing reliable conclusion about household behavior. Results from our preferred mixture specification suggest that relatively modest participation cost can rationalize prevalence of non-participation in the market for risky assets in Australia, with the per-period cost in the range of 500-700 dollars explaining 90% of non-participation in our sample.

The rest of the paper is organized as follows. The data used in the paper is described in the next section. Section 3 presents the sample selection models with normal and mixture of normals disturbances and develops the MCMC algorithm for the Bayesian inference in these models. Section 4 derives expressions for marginal effects. Section 5 discusses the empirical results, section 6 presents and implements the method for calculation of the cost of stock market participation and section 7 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.

The paper focuses on the portfolio decisions of the working age households. This restriction is needed to properly account for the influence of retirement assets (superannuation) in household financial decisions. Personal retirement assets in Australia are generally not accessible before the age of 65, and thus should be excluded from the definition of liquid financial wealth for workers under 65 years old. On the other hand, for households with heads over 65 years old, superannuation assets are properly classified as liquid wealth. To avoid complications created by the changing nature of superannuation funds we restrict our sample by eliminating households in which head is over 65 years old.

The data at our disposal provides information on the value of the shares 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 the financial wealth.

Asset variables which control for background risk faced by a household include retirement wealth (superannuation) and business equity. Superannuation assets in Australia can be invested in stock market, property or safe assets. Because we don't observe the composition of the superannuation portfolio held by a given households and its correlation with the stock market returns it is difficult to predict a priori how superannuation assets should affect participation and allocation decisions. Substantial holdings of business equity, on the other hand, are generally found to reduce household exposure to stock market risk. Following Guiso et al. (1996) and Hochguertel (2003) we use subjective expectations of becoming unemployed as measure of background risk stemming from the uncertainty about labor income. In the previous studies similar measures 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 this paper models decisions whether to invest in risky assets and how much to invest conditional on the set of household's demographic characteristics, employment status, income, risk attitudes, planning horizon and measures of background risk. 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. 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.

In the empirical implementation we will use flexible functional forms to accommodate potential non-linear effects of age, net worth, income and superannuation assets on participation and allocation decisions. After eliminating households for which data for at least one variable used in the analysis is missing the final sample consists of 4110 households, of which 1698 or 41% hold some part of their wealth in risky assets. Variable definitions and summary statistics are given in Table 1.

Table 1: Variable definition and summary statistics (number of observations: 4110)

Variable	Definition	Mean	SD
Financial variables			
equity	=1 if holds public equity	0.41	0.49
fw	Financial Wealth/10000	63.4	204.6
share	share of risky assets in financial wealth	0.53	0.32
nw	Net Worth/10000	38.9	61.0
income	Household Income/10000	5.0	5.0
super	Superannuation/10000	8.7	15.7
bizeq	Business Equity/10000	4.0	26.6
Household Characteristics			
age	age of household head (HH)	43.1	11.7
edub	=1 if HH has bachelor qualification	0.23	0.42
edud	=1 if HH has advanced diploma	0.37	0.48
edus	=1 if HH is high school graduate	0.11	0.32
olf	=1 if HH is out of labor force	0.18	0.38
unemployed	=1 if HH is unemployed	0.03	0.18
liq1	=1 if household is experiencing liquidity constraints	0.23	0.42
prob	=1 if HH is uncertain about the future of his/her a job	0.30	0.46
lone	=1 if lone parent	0.11	0.32
couple	=1 if couple	0.63	0.48
children04	number of children under age of 4	0.21	0.52
children415	number of children between ages of 4 and 15	0.46	0.86
horizon1	=1 if planning horizon is next 2-4 years	0.13	0.33
horizon2	=1 if planning horizon is next 5-10 years	0.23	0.42
risk	=1 if HH is willing to take high risks	0.11	0.31
nesb	=1 if HH comes form non-English speaking background	0.12	0.32
urban	=1 if household resides in urban area	0.61	0.49

3 Normal and Mixture of Normals Sample Selection Models

This section defines the sample selection models with normal and mixture of normals disturbances which will be used to study portfolio choices of Australian households. It also describes the Markov Chain Monte Carlo algorithms which is used to conduct inference in the two models.

Let I_i^* denote the latent utility that an individual derives from stock market participation, let S_i^* denote the *potential* proportion of wealth he would be willing to invest in the stock-market. The model for the vector $[I_i^*, S_i^*]'$ can be specified as follows:

$$
\begin{aligned}
S_i^* &= \boldsymbol{\alpha}'\mathbf{x}_i + \varepsilon_{1i} && (1)\\
I_i^* &= \boldsymbol{\beta}'\mathbf{z}_i + \varepsilon_{2i} && (2)
\end{aligned}
$$

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 (instruments). Without loss of generality we assume that $\mathbf{z}_i = [\mathbf{x}_i', \mathbf{z}_{2i}']'$ where $\mathbf{z}_{2i}'$ is a vector of covariates not included in the share equation (1). In the normal sample selection model the vector of disturbances $\boldsymbol{\varepsilon}_i = [\varepsilon_{1i}, \varepsilon_{2i}]'$ has a bivariate normal distribution with zero mean and variance-covariance matrix Σ:

$$
\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.
$$

In the sample selection model with mixture of normals disturbances the vector $\boldsymbol{\varepsilon}_i$ follows a discrete mixture of bivariate normal distributions:

$$
f(\boldsymbol{\varepsilon}_i|\boldsymbol{\theta}) = \sum_{j=1}^{m} \pi_j \phi_2(\boldsymbol{\varepsilon}_i|\boldsymbol{\mu}_j, \Sigma_j),
$$

where $\boldsymbol{\theta}$ denotes the vector of parameters, $\phi_2(.|a, B)$ denotes probability density function of a bivariate normal distribution with mean a and variance-covariance matrix B, π_j denotes the probability of mixture component j, m denotes the number of components in the mixture, $\Sigma_{j=1}^{m}\pi_j = 1$, and

$$
\boldsymbol{\mu}_j = \begin{pmatrix} \mu_{1j} \\ \mu_{2j} \end{pmatrix},
$$

$$\Sigma_j = \begin{pmatrix} \sigma_{1j}^2 & \sigma_{12j} \\ \sigma_{12j} & \sigma_{2j}^2 \end{pmatrix}$$

for $j = 1, ..., m$. In this setup mixture components have no structural interpretation because component labels are not identifies 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 $I_i = 0$ if $I_i^* < 0$ and $I_i = 1$ if $I_i^* >= 0$. Note that the potential proportion of wealth invested in stocks, S_i^*, is only observed when individual actually participates in the stock market, i.e. when $I_i = 1$. Let S_i denote the actual proportion of wealth invested in the stock market. Then $S_i = S_i^*$ if $I_i = 1$ and $S_i = 0$ otherwise. Using marginal-conditional decomposition the likelihood function of the sample selection model with normal disturbances $L(\boldsymbol{\theta}|\mathbf{Data}, N)$ can be written as:

$$L(\boldsymbol{\theta}|\mathbf{Data}, N) =$$
$$\left(1 - \Phi\left(\frac{\boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma_2^2}}\right)\right)^{1-I_i} \cdot \left\langle \Phi\left(\frac{\boldsymbol{\beta}'\mathbf{z}_i + \sigma_{12}(S_i - \boldsymbol{\alpha}'\mathbf{x}_i)/\sigma_1^2}{\sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2}}\right) \cdot \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(S_i - \boldsymbol{\alpha}'\mathbf{x}_i)^2}{2\sigma_1^2}\right)\right\rangle^{I_i},$$

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

$$L(\boldsymbol{\theta}|\mathbf{Data}, M) = \Pi_{i=1}^N \Sigma_{j=1}^m \pi_j \left(1 - \Phi\left(\frac{\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i}{\sqrt{\sigma_{2j}^2}}\right)\right)^{1-I_i}$$
$$\cdot \left\langle \Phi\left(\frac{\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i + \sigma_{12j}(S_i - \mu_{1j} - \boldsymbol{\alpha}'\mathbf{x}_i)/\sigma_{1j}^2}{\sqrt{\sigma_{2j}^2 - \sigma_{12j}^2/\sigma_{1j}^2}}\right) \cdot \frac{1}{\sqrt{2\pi}\sigma_{1j}} \exp\left(-\frac{(S_i - \mu_{1j} - \boldsymbol{\alpha}'\mathbf{x}_i)^2}{2\sigma_{1j}^2}\right)\right\rangle^{I_i},$$

where $\Phi(a)$ denotes standard normal cdf evaluated at a. It is easy to see that the variance σ_2^2 in the sample selection model is not identified separately from σ_{12} and $\boldsymbol{\beta}$ in the sense that if we multiply σ_2, σ_{12} and $\boldsymbol{\beta}$ by the same constant, the likelihood function would not change. Similarly, in the sample selection model with mixture of normals disturbances the variance terms σ_{2j}^2 are not identified separately from σ_{12j}, μ_{2j} and $\boldsymbol{\beta}$. The identification in the normal sample selection model is achieved by the normalization $\sigma_2^2 = 1$. In the similar fashion, the identification on the sample selection model with mixture of normals disturbances is achieved by the normalization $\sigma_{2j}^2 = 1$ for $j = 1, .., m$.

The Bayesian inference in the sample selection 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^* and the potential proportion of wealth invested in the stock market S_i^*, for $i = 1, ..., N$. The joint probability density function of the augmented data $\mathbf{y}^* = [S_1^*, ..., S_N^*, I_1^*, ..., I_N^*]'$, $\mathbf{S} = [S_1, ..., S_N]'$ 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:

$$
\begin{aligned}
& P(\mathbf{y}^*, \mathbf{I}, \mathbf{S}|\mathbf{Z}, \boldsymbol{\theta}) = |H|^{N/2} \exp\left(-(\mathbf{y}^* - \mathbf{W}\boldsymbol{\gamma})'(H \otimes D_N)(\mathbf{y}^* - \mathbf{W}\boldsymbol{\gamma})/2\right) \\
\cdot \quad & \Pi_{i=1}^N (\iota(S_i = S_i^*)\iota(I_i = 1)\iota(I_i^* \geq 0) + \iota(S_i = 0)\iota(I_i = 0)\iota(I_i^* < 0)),
\end{aligned} \tag{3}
$$

where $H = \Sigma^{-1}$, $\boldsymbol{\gamma} = [\boldsymbol{\alpha}', \boldsymbol{\beta}]'$,

$$
\mathbf{W} = \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z} \end{bmatrix},
$$

$\mathbf{X} = [\mathbf{x}_1', ..., \mathbf{x}_N']$ and D_N is an identity matrix of size $N \times N$.

The model must be completed with the specification of prior distribution of parameters $P(\boldsymbol{\theta})$. The collection of parameters $\boldsymbol{\theta}$ in the normal sample selection model consists of H and $\boldsymbol{\gamma}$. We specify conjugate prior distributions for $\boldsymbol{\gamma}$ and H, and assume that $\boldsymbol{\gamma}$ and H are independent in the prior:

$$
P(\boldsymbol{\theta}) = P(\boldsymbol{\gamma}) \cdot P(H), \tag{4}
$$

where

- $\boldsymbol{\gamma} \sim N(\underline{\boldsymbol{\gamma}}, \underline{H}_\gamma)$
- $H \sim Wishart(\underline{S}, \nu) \cdot \iota(\sigma_2^2 = 1)$.

Note, that the prior distribution of H is truncated so that $\sigma_2^2 = 1$ to satisfy the normalization constraints.

The joint posterior distribution of $\boldsymbol{\theta}$ is proportional to the product of (3) and (4). To approximate the posterior distribution of $\boldsymbol{\theta}$ we construct the Gibbs sampling algorithm which iterates between the following three blocks:

1. Sample $[S_i^*, I_i^*]'|S_i, I_i, \mathbf{z}_i, \boldsymbol{\theta}$ for $i = 1, ..., N$. When $I_i = 1$, the conditional posterior distribution of S_i^* is degenerate at S_i, so we set $S_i^* = S_i$ and draw $I_i^*|(S_i^*, S_i, I_i, \mathbf{z}_i, \boldsymbol{\theta}) \sim N(\boldsymbol{\beta}'\mathbf{z}_i + \frac{\sigma_{12}}{\sigma_1^2}(S_i^* - \boldsymbol{\alpha}'\mathbf{x}_i), 1 - \frac{\sigma_{12}^2}{\sigma_1^2})$ truncated to $I_i^* > 0$;

When $I_i = 0$, $[I_i^*, S_i^*]'|S_i, I_i, \mathbf{z}_i, \boldsymbol{\theta}) \sim N\left(\begin{array}{c} \boldsymbol{\alpha}'\mathbf{z}_i \\ \boldsymbol{\beta}'\mathbf{z}_i \end{array}, \Sigma\right)$ truncated to $I_i^* < 0$. To sample from this joint distribution we first draw from the marginal-conditional posterior $I_i^*|(S_i, I_i, \mathbf{z}_i, \boldsymbol{\theta}) \sim N(\boldsymbol{\beta}'\mathbf{z}_i, 1)$ truncated to $I_i^* < 0$, and then draw from the conditional posterior $S_i^*|(S_i, I_i^*, I_i, \mathbf{z}_i, s_i, \boldsymbol{\theta}) \sim N(\boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12}(I_i^* - \boldsymbol{\beta}'\mathbf{z}_i), \sigma_{11} - \sigma_{12}^2)$. This algorithm for drawing from the joint posterior conditional distribution $[S_i^*, I_i^*]'|S_i, I_i, \mathbf{z}_i, \boldsymbol{\theta}$ is due to van Hasselt (2007).

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

3. $H|(\mathbf{y}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\gamma}) \sim Wishart(\overline{S}, \overline{\nu}) \cdot \iota(\sigma_2^2 = 1)$ where

$$\overline{S} = \underline{S} + \sum_{i=1}^{N} \left[\begin{array}{ll} (S_i^* - \boldsymbol{\alpha}'\mathbf{x}_i)^2 & (S_i^* - \boldsymbol{\alpha}'\mathbf{x}_i)(I_i^* - \boldsymbol{\beta}'\mathbf{z}_i) \\ (S_i^* - \boldsymbol{\alpha}'\mathbf{x}_i)(I_i^* - \boldsymbol{\beta}'\mathbf{z}_i) & (I_i^* - \boldsymbol{\beta}'\mathbf{z}_i)^2 \end{array} \right]$$

and $\overline{\nu} = \underline{\nu} + N$. The draws from this truncated Wishart distribution are obtained using the algorithm proposed in Nobile (2000).

Similar to the normal model, Bayesian inference in the sample selection 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^*, the potential proportion of wealth invested in the stock market S_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 $P(s_i = j|\mathbf{x}_i, \mathbf{x}_{i1}, x_{i2}, \boldsymbol{\theta}) = \pi_j$ for $j = 1, ..., m$. The distribution of the disturbances $\boldsymbol{\varepsilon}_i$ conditional on the latent indicator of mixture component s_i is normal:

$$\boldsymbol{\varepsilon}_i|(s_i = j, \boldsymbol{\theta}) \sim N(\boldsymbol{\mu}_j, \Sigma_j).$$

The following notation will be useful for the presentation of the posterior simulation algorithm in the sample selection model with mixture of normals disturbances. Define $\mathbf{C} = [\mathbf{c}_1, ..., \mathbf{c}_N]' = [\iota(s_i = j)]$, 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 $\mathbf{W} = \left[\begin{array}{cc} \mathbf{W}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2 \end{array}\right]$ where $\mathbf{W}_1 = [\mathbf{C}, \mathbf{X}]$ and $\mathbf{W}_2 = [\mathbf{C}, \mathbf{Z}]$, and define $\boldsymbol{\gamma} = [\mu_{11}, ..., \mu_{1m}, \boldsymbol{\alpha}', \mu_{21}, ..., \mu_{2m}, \boldsymbol{\beta}']'$. Then the joint probability density function of the augmented data $\mathbf{s} = [s_1, ..., s_N]'$, $\mathbf{y}^* = [S_1^*, ..., S_N^*, I_1^*, ..., I_N^*]'$, $\mathbf{S} = [S_1, ..., S_N]'$ and $\mathbf{I} = [I_1, ..., I_N]'$ conditional on exogenous variables $\mathbf{Z}$ and the vector of parameters $\boldsymbol{\theta}$ can be written:

$$P(\mathbf{y}^*, \mathbf{I}, \mathbf{S}, \mathbf{s}|\mathbf{Z}, \boldsymbol{\theta}) = \prod_{j=1}^{m} \pi_j^{N_j} |H_j|^{N_j/2} \exp\left(-(\mathbf{y}^* - \mathbf{W}\boldsymbol{\gamma})'\mathbf{H}(\mathbf{y}^* - \mathbf{W}\boldsymbol{\gamma})/2\right)$$
$$\cdot \quad \Pi_{i=1}^{N}(\iota(S_i = S_i^*)\iota(I_i = 1)\iota(I_i^* \geq 0) + \iota(S_i = 0)\iota(I_i = 0)\iota(I_i^* < 0)), \tag{5}$$

where N_j is the number of observations such that $s_i = j$,

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$$

and $H_{kl} = diag(\Sigma_{s_1}^{kl}, ..., \Sigma_{s_N}^{kl})$ where Σ_j^{kl} is kl'th element of Σ_j^{-1} for $k = 1, 2$, $l = 1, 2$, $j = 1, ..., m$.

The model must be completed with the specification of prior distribution of parameters $P(\boldsymbol{\theta})$. The collection of parameters $\boldsymbol{\theta}$ consists of $\boldsymbol{\gamma}, H_1, ..., H_m$ and $\boldsymbol{\pi}$, where $H_j \equiv \Sigma_j^{-1}$ for $j = 1, ..., m$. We specify conjugate prior distributions for $\boldsymbol{\alpha}, H_1, ..., H_m, \boldsymbol{\pi}$ and assume that $\boldsymbol{\alpha}, H_1, ..., H_m, \boldsymbol{\pi}$ are independent in the prior:

$$P(\boldsymbol{\theta}) = P(\boldsymbol{\gamma}) \cdot \prod_{j=1}^{m} P(H_j) \cdot P(\boldsymbol{\pi}), \tag{6}$$

where

- $\boldsymbol{\gamma} \sim N(\underline{\boldsymbol{\gamma}}, \underline{H}_\gamma)$
- $H_j \sim Wishart(\underline{S}_j, \nu_j) \cdot \iota(\sigma_{2j}^2 = 1)$ for $j = 1, ..., m$
- $\boldsymbol{\pi} \sim Dirichlet(r)$

Note, that the prior distribution of H_j is truncated so that $\sigma_{2j}^2 = 1$ to satisfy the normalization constraints.

The joint posterior distribution of $\boldsymbol{\theta}$ is proportional to the product of (5) and (6). To approximate the posterior distribution of $\boldsymbol{\theta}$ we construct the Gibbs sampler which iterates between the following five blocks:

1. Sample $[S_i^*, I_i^*]'|S_i, I_i, \mathbf{z}_i, s_i, \boldsymbol{\theta}$ for $i = 1, ..., N$. When $I_i = 1$, set $S_i^* = S_i$ and draw $I_i^*|(S_i^*, S_i, I_i, \mathbf{z}_i, s_i, \boldsymbol{\theta}) \sim N(\mu_{2s_i} + \boldsymbol{\beta}'\mathbf{z}_i + \frac{\sigma_{12s_i}}{\sigma_{1s_i}^2}(S_i^* - \mu_{1s_i} - \boldsymbol{\alpha}'\mathbf{x}_i), 1 - \frac{\sigma_{12s_i}^2}{\sigma_{1s_i}^2})$ truncated to $I_i^* > 0$;

   When $I_i = 0$, draw $I_i^*|(S_i, I_i, \mathbf{z}_i, s_i, \boldsymbol{\theta}) \sim N(\mu_{2s_i} + \boldsymbol{\beta}'\mathbf{z}_i, 1)$ truncated to $I_i^* < 0$ and then draw $S_i^*|(S_i, I_i^*, I_i, \mathbf{z}_i, s_i, \boldsymbol{\theta}) \sim N(\mu_{1s_i} + \boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12s_i}(I_i^* - \mu_{2s_i} - \boldsymbol{\beta}'\mathbf{z}_i), \sigma_{11s_i} - \sigma_{12s_i}^2)$;

2. $P(s_i = j|S_i^*, S_i, I_i^*, I_i, \mathbf{z}_i, \boldsymbol{\theta}) \propto \pi_j \phi(\begin{bmatrix} S_i^* \\ I_i^* \end{bmatrix}; \boldsymbol{\mu}_j + \begin{bmatrix} \boldsymbol{\alpha}'\mathbf{x}_i \\ \boldsymbol{\beta}'\mathbf{z}_i \end{bmatrix}, \Sigma_j)$ where $\phi(a, B)$ denotes probability density function of a bivariate normal distribution with mean a and variance-covariance matrix B.

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

4. $H_j|(\mathbf{y}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \mathbf{s}, \boldsymbol{\theta}_{-H_j}) \sim Wishart(\overline{S}_j, \overline{\nu}_j) \cdot \iota(\sigma_{2j}^2 = 1)$ where

$$\overline{S}_j = \underline{S}_j + \sum_{i=1}^{N} \iota(s_i = j) \begin{bmatrix} (S_i^* - \mu_{1j_i} - \boldsymbol{\alpha}'\mathbf{x}_i)^2 & (S_i^* - \mu_{1j_i} - \boldsymbol{\alpha}'\mathbf{x}_i)(I_i^* - \mu_{2j_i} - \boldsymbol{\beta}'\mathbf{z}_i) \\ (S_i^* - \mu_{1j_i} - \boldsymbol{\alpha}'\mathbf{x}_i)(I_i^* - \mu_{2j_i} - \boldsymbol{\beta}'\mathbf{z}_i) & (I_i^* - \mu_{2j_i} - \boldsymbol{\beta}'\mathbf{z}_i)^2 \end{bmatrix}$$

and $\overline{\nu}_j = \underline{\nu}_j + N_j$. The draws from this truncated Wishart distribution are obtained using the algorithm proposed in Nobile (2000).

6. $\boldsymbol{\pi}|(\mathbf{y}^*, \mathbf{S}, \mathbf{I}, \mathbf{Z}, \boldsymbol{\theta}_{-\boldsymbol{\pi}}) \sim Dirichlet(r + N_1, ..., r + N_m)$.

The Matlab codes for these two algorithms have passed the joint distribution test suggested in Geweke (2004). Next we discuss posterior distributions of marginal effects in the sample selection models with normal and mixture of normals disturbances.

4 Marginal Effects

Results of the normal and mixture of normals sample selection models can be interpreted by computing 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}:

$$MEP_{z_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv \frac{\partial Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A)}{\partial z_{ki}}, \tag{7}$$

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 in the conditioning set indicates model for which the effect is computed, i.e. $A = N$ for sample selection model with normal disturbances and $A = M$ for sample selection model with mixture of normals disturbances.

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}:

$$MEP_{z_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv Prob(I_i^* > 0|\mathbf{z}_{-z_{ki},i}, z_{ki} = a+1, \boldsymbol{\theta}, A) - Prob(I_i^* > 0|\mathbf{z}_{-z_{ki},i}, z_{ki} = a, \boldsymbol{\theta}, A), \tag{8}$$

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 expectation of 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}:

$$MESC_{x_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv \frac{\partial E(S_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, A)}{\partial x_{ki}}. \tag{9}$$

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

$$MESC_{x_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv E(S_i^*|I_i^* > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a+1, \boldsymbol{\theta}, A) - E(S_i^*|I_i^* > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a, \boldsymbol{\theta}, A). \tag{10}$$

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

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

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 is equal to the potential fraction S_i^* for individuals who participate in the stock market, the expectation $E(S_i|\mathbf{z}_i, \boldsymbol{\theta}, A)$ reduces to:

$$E(S_i|\mathbf{z}_i, \boldsymbol{\theta}, A) = E(S_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, A) \cdot Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, A). \tag{11}$$

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

$$MESU_{x^c_{ki}}|\mathbf{z}_i, \boldsymbol{\theta}, A \equiv \frac{\partial E(S^*_i|I^*_i > 0, \mathbf{z}_i, \boldsymbol{\theta}, A) Prob(I^*_i > 0|\mathbf{z}_i, \boldsymbol{\theta}, A)}{\partial x_{ki}}. \tag{12}$$

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

$$\begin{aligned} &MESU_{x^d_{ki}}|\mathbf{z}_i, \boldsymbol{\theta},\mathrm{A} \equiv \\ &E(S^*_i|I^*_i > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a+1, \boldsymbol{\theta}, A) Prob(I^*_i > 0|\mathbf{z}_{-x_{ki},i}, x_{ki} = a+1, \boldsymbol{\theta}, A) - \\ &E(S^*_i|I^*_i > 0, \mathbf{z}_{-x_{ki},i}, x_{ki} = a, \boldsymbol{\theta}, A) Prob(I^*_i > 0|\mathbf{z}_{-x_{ki},i}, x_{ki} = a, \boldsymbol{\theta}, A). \end{aligned} \tag{13}$$

Now we will derive the expressions for the marginal effects (7)-(13) in sample selection models with normal and mixture of normals disturbances. In the sample selection model with normal disturbances the conditional probability of stock market participation $Prob(I^*_i > 0|\mathbf{z}_i, \boldsymbol{\theta}, N)$ is equal to the standard normal cdf evaluated at marginal expectation of $I^*_i|\mathbf{z}_i, \boldsymbol{\theta}, N$ $\boldsymbol{\beta}'\mathbf{z}_i$:

$$Prob(I^*_i > 0|\mathbf{z}_i, \boldsymbol{\theta}, N) = \Phi(\boldsymbol{\beta}'\mathbf{z}_i) \tag{14}$$

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:

$$MEP_{z^c_{ki}}|\mathbf{z}_i, \boldsymbol{\theta}, N = \frac{\partial \Phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\partial z_{ki}} = \beta_k \cdot \phi(\boldsymbol{\beta}'\mathbf{z}_i), \tag{15}$$

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:

$$\begin{aligned} MEP_{z^d_{ki}}|\mathbf{z}_i, \boldsymbol{\theta}, N &= \Phi(\beta_k(z_{ki}+1) + \boldsymbol{\beta}_{-\beta_k}\mathbf{z}_{-z_{ki}}) - \Phi(\boldsymbol{\beta}'\mathbf{z}_i) \\ &= \Phi(\beta_k + \boldsymbol{\beta}'\mathbf{z}_i) - \Phi(\boldsymbol{\beta}'\mathbf{z}_i). \end{aligned} \tag{16}$$

To obtain expressions for marginal effects of covariates on fraction of wealth invested in shares we first derive the expected value of fraction of wealth invested in shares conditional on participation $E(S^*_i|I^*_i > 0, \mathbf{z}_i, \boldsymbol{\theta}, N)$. The expression for this conditional expectation can be obtained using standard results about moments of incidentally truncated bivariate normal distribution (e.g. Greene, Theorem 22.5):

$$E(S_i^*|I_i^*>0,\mathbf{z}_i,\boldsymbol{\theta},N)=\boldsymbol{\alpha}'\mathbf{x}_i+\sigma_{12}\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)}, \tag{17}$$

where $\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)}$ is the inverse Mills ratio. Then the marginal effect of a continuous x_{ki} on this conditional expectation is given by:

$$MES_{x_{ki}^c}|\mathbf{z}_i,\boldsymbol{\theta},N=\frac{\partial E(S_i^*|I_i^*>0,\mathbf{z}_i,\boldsymbol{\theta})}{\partial x_{ki}}=\alpha_k-\beta_k\cdot\sigma_{12}\cdot\left(\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)}\left(\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)}+\boldsymbol{\beta}'\mathbf{z}_i\right)\right). \tag{18}$$

The marginal effect of a discrete x_{ki} on the conditional expectation $E(S_i^*|I_i^*>0,\mathbf{z}_i,\boldsymbol{\theta},N)$ is given by:

$$\begin{aligned}MES_{x_{ki}^d}|\mathbf{z}_i,\boldsymbol{\theta},N=\alpha_k+\boldsymbol{\alpha}'\mathbf{x}_i+\sigma_{12}\frac{\phi(\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)}-\boldsymbol{\alpha}'\mathbf{x}_i-\sigma_{12}\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)}=\\ \alpha_k+\sigma_{12}\left(\frac{\phi(\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)}-\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)}\right).\end{aligned} \tag{19}$$

Finally, the unconditional expectation of the observed fraction of wealth invested in shares in the normal sample selection model $E(S_i|\mathbf{z}_i,\boldsymbol{\theta},N)$ is given by

$$E(S_i|\mathbf{z}_i,\boldsymbol{\theta},N)=(\boldsymbol{\alpha}'\mathbf{x}_i+\sigma_{12}\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)})\cdot\Phi(\boldsymbol{\beta}'\mathbf{z}_i). \tag{20}$$

Then the marginal effect of a continuous x_{ki} on this expected value is given by:

$$\begin{aligned}MESU_{x_{ki}^c}|\mathbf{z}_i,\boldsymbol{\theta},N=\frac{\partial E(S_i|\mathbf{z}_i,\boldsymbol{\theta})}{\partial x_{ki}}=\beta_k\cdot\phi(\boldsymbol{\beta}'\mathbf{z}_i)\cdot(\mathbf{x}_i'\boldsymbol{\alpha}+\sigma_{12}\frac{\phi(\mathbf{z}_i'\boldsymbol{\beta})}{\Phi(\mathbf{z}_i'\boldsymbol{\beta})})+\\ \Phi(\mathbf{z}_i'\boldsymbol{\beta})\cdot(\alpha_k-\beta_k\cdot\sigma_{12}\cdot\left(\frac{\phi(\mathbf{z}_i'\boldsymbol{\beta})}{\Phi(\mathbf{z}_i'\boldsymbol{\beta})}\left(\frac{\phi(\mathbf{z}_i'\boldsymbol{\beta})}{\Phi(\mathbf{z}_i'\boldsymbol{\beta})}+\mathbf{z}_i'\boldsymbol{\beta}\right)\right)),\end{aligned} \tag{21}$$

and the marginal effect of a discrete x_{ki} on this expected value is given by:

$$\begin{aligned}MESU_{x_{ki}^d}|\mathbf{z}_i,\boldsymbol{\theta},N&=(\alpha_k+\boldsymbol{\alpha}'\mathbf{x}_i+\sigma_{12}\frac{\phi(\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)})\cdot\Phi(\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)-\\ &\quad(\boldsymbol{\alpha}'\mathbf{x}_i+\sigma_{12}\frac{\phi(\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\boldsymbol{\beta}'\mathbf{z}_i)})\cdot\Phi(\boldsymbol{\beta}'\mathbf{z}_i).\end{aligned} \tag{22}$$

Now we will derive expressions for the marginal effects (7)-(13) in the sample selection model with mixture of normals disturbances. In this model the probability of stock market participation of individual i $Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, M)$ is equal to the sum of mixture-component specific probabilities of participation weighted by the marginal probabilities of these components:

$$Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, M) = \sum_{j=1}^{m} \pi_j Prob(I_i^* > 0|\mathbf{z}_i, \boldsymbol{\theta}, s_i = j, M) = \sum_{j=1}^{m} \pi_j \Phi(\mu_{2i} + \boldsymbol{\beta}'\mathbf{z}_i). \tag{23}$$

Hence, the marginal effect of z_{ki} on probability of stock market participation of individual i is equal to the mixture-component specific marginal effects weighted by the marginal probabilities of these components. In particular, the marginal effect of a continuous z_{ki} is given by:

$$MEP_{z_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, M = \sum_{i=1}^{m} \pi_i \frac{\partial \Phi(\mu_{2i} + \boldsymbol{\beta}'\mathbf{z}_i)}{\partial z_{ki}} = \beta_k \cdot \sum_{i=1}^{m} \pi_i \cdot \phi(\mu_{2i} + \boldsymbol{\beta}'\mathbf{z}_i), \tag{24}$$

and that of a discrete z_{ki} is given by:

$$MEP_{z_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, M = \sum_{j=1}^{m} \pi_j (\Phi(\mu_{2j} + \beta_k + \boldsymbol{\beta}'\mathbf{z}_i) - \Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)). \tag{25}$$

To obtain the expression for the marginal effects of x_{ki} on expected fraction of wealth invested in shares conditional on participation $E(S_i^*|I_i^* > 0, \boldsymbol{\theta}, \mathbf{z}_i, M)$ we first derive the expression for this conditional expectation as follows:

$$\begin{aligned}
& E(S_i^*|I_i^* > 0, \boldsymbol{\theta}, \mathbf{z}_i, M) = \int \int_0^{\infty} S_i^* \cdot p(S_i^*, I_i^*|I_i^* > 0, \boldsymbol{\theta}, \mathbf{z}_i, M) dI_i^* dS_i^* \\
= \; & \int \int_0^{\infty} S_i^* \frac{\sum_{j=1}^{m} \pi_j p(S_i^*, I_i^*|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)}{\sum_{j=1}^{m} \pi_j Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)} dI_i^* dS_i^* \\
= \; & \frac{\sum_{j=1}^{m} \pi_j Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i) \cdot \int \int_0^{\infty} S_i^* \frac{p(S_i^*, I_i^*|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)}{Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)} dI_i^* dS_i^*}{\sum_{j=1}^{m} \pi_j Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)} \\
= \; & \frac{\sum_{j=1}^{m} \pi_j Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i) \cdot E(S_i^*|I_i^* > 0, \boldsymbol{\theta}, s_i = j, \mathbf{z}_i)}{\sum_{i=1}^{m} \pi_j Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)} \\
= \; & \frac{\sum_{j=1}^{m} \pi_j \Phi(\mu_{2j} + \mathbf{z}_i'\boldsymbol{\beta}) \cdot \left(\mu_{1j} + \mathbf{x}_i'\boldsymbol{\alpha} + \sigma_{12j} \frac{\phi(\mu_{2j} + \mathbf{z}_i'\boldsymbol{\beta})}{\Phi(\mu_{2j} + \mathbf{z}_i'\boldsymbol{\beta})} \right)}{\sum_{j=1}^{m} \pi_j \Phi(\mu_{2j} + \mathbf{z}_i'\boldsymbol{\beta})}.
\end{aligned} \tag{26}$$

Hence, the conditional expectation $E(S_i^*|I_i^* > 0, \boldsymbol{\theta}, \mathbf{z}_i, M)$ is equal to the weighted average of the mixture-component specific conditional expectations $E(S_i^*|I_i^* > 0, \boldsymbol{\theta}, s_i = j, \mathbf{z}_i)$ with weights equal to the probabilities of mixture components conditional on participation $p(s_i = j|I_i^* > 0, \boldsymbol{\theta}, \mathbf{z}_i) = \frac{p(I_i^*>0|\mathbf{z}_i,\boldsymbol{\theta},s_i=j)p(s_i=j|\mathbf{z}_i,\boldsymbol{\theta})}{p(I_i^*>0|\mathbf{z}_i,\boldsymbol{\theta})} = \frac{\pi_j\Phi(\mu_{2j}+\mathbf{z}_i'\boldsymbol{\beta})}{\sum_{j=1}^m \pi_j\Phi(\mu_{2j}+\mathbf{z}_i'\boldsymbol{\beta})}$. Then the marginal effect of a continuous variable x_{ki} on the expected fraction of wealth invested in shares among those who participate in the stock market is given by:

$$\begin{aligned} & MES_{x_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, M = \frac{\partial E(S_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta})}{\partial x_{ki}} \\ = & \frac{(\sum_{j=1}^m \pi_j(\beta_k\phi_j E_j + \Phi_j E'_{kj}))\cdot(\sum_{j=1}^m \pi_j\Phi_j) - (\beta_k\sum_{j=1}^m \pi_j\phi_j)\cdot(\sum_{j=1}^m \pi_j\Phi_j E_j)}{(\sum_{j=1}^m \pi_j\Phi_j)^2}, \end{aligned} \tag{27}$$

where $\phi_j \equiv \phi(\mu_{2j} + \mathbf{z}'\boldsymbol{\beta})$, $\Phi_j \equiv \Phi(\mu_{2j} + \mathbf{z}'\boldsymbol{\beta})$,

$$\begin{aligned} & E_j \equiv E(S_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, s_i = j) \\ = \; & \mu_{1j} + \mathbf{x}'\boldsymbol{\alpha} + \sigma_{12j}\frac{\phi(\mu_{2j} + \mathbf{z}'\boldsymbol{\beta})}{\Phi(\mu_{2j} + \mathbf{z}'\boldsymbol{\beta})}, \\ & E'_{kj} \equiv \frac{\partial E(S_i^*|I_i^* > 0, \mathbf{z}_i, \boldsymbol{\theta}, s_i = j)}{\partial x_{ki}} \\ = \; & \alpha_k - \beta_k\cdot\sigma_{12j}\cdot\left(\frac{\phi(\mu_{2j} + \mathbf{z}'\boldsymbol{\beta})}{\Phi(\mu_{2j} + \mathbf{z}'\boldsymbol{\beta})}\left(\frac{\phi(\mu_{2i} + \mathbf{z}'\boldsymbol{\beta})}{\Phi(\mu_{2i} + \mathbf{z}'\boldsymbol{\beta})} + \mu_{2i} + \mathbf{z}'\boldsymbol{\beta}\right)\right). \end{aligned}$$

The marginal effect of a discrete x_{ki} on the expected fraction of wealth invested in stock market conditional on participation is given by:

$$\begin{aligned} & MES_{x_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta}, M = \\ = & \frac{\sum_{j=1}^m \pi_j\Phi(\mu_{2j} + \beta_k + \boldsymbol{\beta}'\mathbf{z}_i)\cdot\left(\mu_{1j} + \alpha_k + \boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12j}\frac{\phi(\mu_{2j}+\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2j}+\beta_k+\boldsymbol{\beta}'\mathbf{z}_i)}\right)}{\sum_{j=1}^m \pi_j\Phi(\mu_{2j} + \beta_k + \boldsymbol{\beta}'\mathbf{z}_i)} - \\ & \frac{\sum_{j=1}^m \pi_j\Phi(\mu_{2j} + \mathbf{z}_i'\boldsymbol{\beta})\cdot\left(\mu_{1j} + \boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12j}\frac{\phi(\mu_{2j}+\boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2j}+\boldsymbol{\beta}'\mathbf{z}_i)}\right)}{\sum_{j=1}^m \pi_j\Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}. \end{aligned} \tag{28}$$

The unconditional expectation of the observed fraction of wealth invested in shares in the sample selection model with mixture of normals disturbances $E(S_i|\boldsymbol{\theta}, \mathbf{z}_i, M)$ can be expressed:

$$\begin{aligned} & E(S_i|\boldsymbol{\theta}, \mathbf{z}_i, M) = Prob(I_i^* > 0|\boldsymbol{\theta}, \mathbf{z}_i)\cdot\int\int_0^\infty S_i^*\cdot p(S_i^*, I_i^*|I_i^* > 0, \boldsymbol{\theta}, \mathbf{z}_i)dI_i^* dS_i^* \\ = & \int\int_0^\infty S_i^*\sum_{j=1}^m \pi_j p(S_i^*, I_i^*|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)dI_i^* dS_i^* \end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{m} \pi_j Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i) \cdot \int\int_0^{\infty} S_i^* \frac{p(S_i^*, I_i^*|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)}{Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i)} dI_i^* dS_i^* \\
&= \sum_{j=1}^{m} \pi_j Prob(I_i^* > 0|\boldsymbol{\theta}, s_i = j, \mathbf{z}_i) \cdot E(S_i^*|I_i^* > 0, \boldsymbol{\theta}, s_i = j, \mathbf{z}_i) \\
&= \sum_{j=1}^{m} \pi_j \Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i) \cdot \left(\mu_{1j} + \boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12j} \frac{\phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}\right)
\end{aligned}
$$

Then the marginal effect of a continuous x_{ki} on this unconditional expectation is given by:

$$
\begin{aligned}
MESU_{x_{ki}^c}|\mathbf{z}_i, \boldsymbol{\theta}, M &= \sum_{j=1}^{m} \pi_j (\phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i) \left(\mu_{1j} + \boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12j} \frac{\phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}\right) \\
&+ \Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)(\alpha_k - \beta_k \cdot \sigma_{12j} \cdot \left(\frac{\phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)} \left(\frac{\phi(\mu_{2i} + \boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2i} + \boldsymbol{\beta}'\mathbf{z}_i)} + \mu_{2i} + \boldsymbol{\beta}'\mathbf{z}_i\right)\right)), \quad (29)
\end{aligned}
$$

and that of a discrete x_{ki} is given by:

$$
\begin{aligned}
MESU_{x_{ki}^d}|\mathbf{z}_i, \boldsymbol{\theta} &= \sum_{j=1}^{m} \pi_j \Phi(\mu_{2j} + \beta_k + \boldsymbol{\beta}'\mathbf{z}_i) \cdot \left(\mu_{1j} + \alpha_k + \boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12j} \frac{\phi(\mu_{2j} + \beta_k + \boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2j} + \beta_k + \boldsymbol{\beta}'\mathbf{z}_i)}\right) - \\
&\sum_{j=1}^{m} \pi_j \Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i) \cdot \left(\mu_{1j} + \boldsymbol{\alpha}'\mathbf{x}_i + \sigma_{12j} \frac{\phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}{\Phi(\mu_{2j} + \boldsymbol{\beta}'\mathbf{z}_i)}\right). \quad (30)
\end{aligned}
$$

The conditional marginal effects in (15)-(30) all depends 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 couple household whose continuous covariates (household head's age, net worth, income, business equity) are equal to their sample medians, with no children younger than 16 years old, not liquidity constrained, living in urban area of New South Wales and whose household head has a job, is not willing to take risks, has advanced diploma as the highest educational qualification, is not uncertain about the future of his job, comes from English-speaking background and has planning horizon less or equal to one year. To obtain the *posterior distribution* of the marginal effects we evaluate expressions (15)-(30) 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:

$$p(MEf_{z_{ki}^h}|\mathbf{z}_i = \overline{\mathbf{z}}, \mathbf{Data}, A) = \int p(MEf_{z_{ki}^h}|\mathbf{z}_i = \overline{\mathbf{z}}, \boldsymbol{\theta}, A)p(\boldsymbol{\theta}|\mathbf{Data})d\boldsymbol{\theta}, \tag{31}$$

where $f = \{P, S, SU\}$, $h = \{c, d\}$ and $\overline{\mathbf{z}}$ denotes the vector of covariates $\mathbf{z}_i$ of a representative Australian 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

The main goal 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? The sample selection model is well suited for answering both of these questions in a unified framework. It allows to model separately decisions whether to invest in the stock market and how much to invest, while simultaneously controlling for potential non-random selection into stock ownership. Employing the mixture of normals sample selection model allows us to relax normality assumption of the standard model and to avoid problems which model misspecification can potentially cause.

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

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

$$\begin{aligned} \underline{\boldsymbol{\alpha}} &= [.53, \mathbf{0}'_{K_x \times 1}]', \\ \underline{\boldsymbol{\beta}} &= [-.23, \mathbf{0}'_{K_z \times 1}]', \end{aligned}$$

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

   $\underline{H}_\alpha = (1/50)\mathbf{I}_{K_x \times K_x}$ 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 vector $\boldsymbol{\varepsilon}_i$, $\underline{S}$ and $\underline{\nu}$ are specified as follows:

$$\begin{aligned} \underline{S} &= \begin{bmatrix} 1.9 & 0 \\ 0 & 19 \end{bmatrix} \\ \underline{\nu} &= 20 \end{aligned}$$

This prior distribution implies independence between ε_{1i} and ε_{1i} and centers variance of $S_i^*|\mathbf{x}_i, \boldsymbol{\theta}$ around sample variance of $S_i|I_i = 1$, and that of $I_i^*|\mathbf{x}_i, \boldsymbol{\theta}$ around sample variance of I_i.

In the mixture of normals sample selection 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 $\boldsymbol{\gamma}$, $\underline{\boldsymbol{\gamma}} = [\underline{\boldsymbol{\mu}}_1', \underline{\boldsymbol{\alpha}}', \underline{\boldsymbol{\mu}}_2', \underline{\boldsymbol{\beta}}']'$ and the precision of this distribution $\underline{H}_\gamma$ are specified as follows:

$$\begin{aligned} \underline{\mu}_1 &= \mathbf{0}_{m\times 1}, \\ \underline{\boldsymbol{\alpha}} &= [.53, \mathbf{0}'_{K_x\times 1}]', \\ \underline{\mu}_2 &= \mathbf{0}_{m\times 1}, \\ \underline{\boldsymbol{\beta}} &= [-.23, \mathbf{0}'_{K_z\times 1}]', \end{aligned}$$

$$\underline{H}_\gamma = \begin{bmatrix} \underline{H}_{\mu_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \underline{H}_\alpha & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \underline{H}_{\mu_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \underline{H}_\beta \end{bmatrix},$$

where $\underline{H}_{\mu_1} = .1\mathbf{I}_{m\times M}$, $\underline{H}_\alpha = (1/50)\mathbf{I}_{K_x\times K_x}$, $\underline{H}_{\mu_2} = \mathbf{I}_{m\times M}$, $\underline{H}_\beta = (1/50)\mathbf{I}_{K_z\times K_z}$. 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 vector $\boldsymbol{\varepsilon}_i$, $\underline{S}_j$ and $\underline{\nu}_j$, $j = 1, ..., m$, are specified as follows:

$$\begin{aligned} \underline{S}_j &= \begin{bmatrix} 1.9 & 0 \\ 0 & 19 \end{bmatrix} \\ \underline{\nu}_j &= 20 \end{aligned}$$

This prior distribution implies independence between ε_{1i} and ε_{1i} conditional on the component of the mixture and centers variance of $S_i^*|\mathbf{x}_i, \boldsymbol{\theta}, s_i = j$ around sample variance of $S_i|I_i = 1$ for $j = 1, ..., m$, and that of $I_i^*|\mathbf{x}_i, \boldsymbol{\theta}, s_i = j$ around sample variance of I_i.

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.

We fit models with up to five components in the mixture of normals sample selection model, assuming that five component mixture should be general enough to accommodate most of the important features of the data. Model selection is based on the comparison log marginal likelihoods which are computed using the method proposed by Chib (1995). As shown in table 2, the data favors the 5 component mixture over all other specifications. The

Table 2: Log Marginal Likelihood Comparison

Model	Log of Marginal Likelihood
Normal	-2780.2
Mixture of 2 Normals	-2757.8
Mixture of 3 Normals	-2748.9
Mixture of 4 Normals	-2705.4
Mixture of 5 Normals	**-2247.4**

normal model is strongly rejected by the data and ranks last in terms of marginal likelihood comparison. In what follows we treat the five component mixture as our preferred model and compare its performance to the normal model.

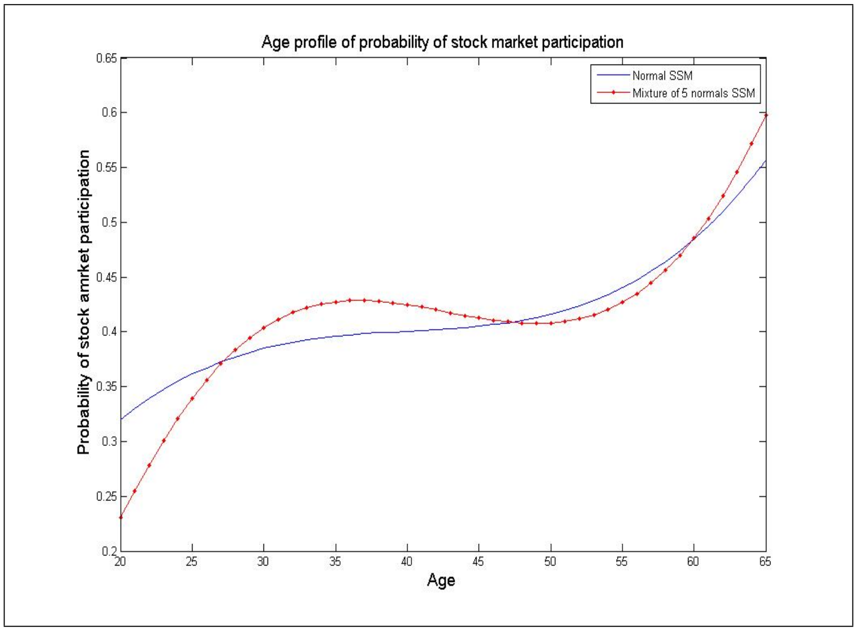
Empirical results for the five component mixture model are presented in tables 3, 5 and 7. Table 3 presents posterior moments of the coefficients and marginal effects in the participation equation. Table 5 contains posterior moments of the coefficients in the share equation and posterior moments of the 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 normal 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, the identification strategy assumes that such variables as non-English speaking background, urban residence and state of residence are assumed to influence only participation decision and therefore can be excluded from the share equation. 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 11% less likely to invest in stocks. We do not find any significant effect of being urban resident on the probability to participate in the stock market. Finally, there exist significant differences in participation rates across states, with residents of Queensland, Tasmania and Western Australia being 6%, 6% and 7.5% less likely to hold stocks respectively.

As can be seen in Table 3, households with higher net worth and income are more likely to invest in stocks, although the effects of these variables is modest, i.e. increase in net worth or income by 10000 dollars increases participation by about 0.01%. Participation is also increasing with the level of education: households headed by persons with 12 years of schooling are 11% more likely to hold public equity compared to those with 9 years of schooling. Interestingly, additional education beyond 12 years does not seem to increase the probability of participation. The median of the age profile of stock market participation for a representative Australian household as implied by the model is presented in Figure 1. The effect of age is found to be non-linear: participation is increasing between ages of 20 and 35, remains relatively constant until the age of 53, and then again is rapidly increasing until the age of 65.

Table 3 also implies that planning horizon and attitude towards risk have strong effects on the 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 20% more likely to hold risky assets. Households with unemployed head are 10% less likely to participate in the stock market, while for those experiencing liquidity constraints this probability is reduced by 13%. The magnitudes of marginal effects of planning horizon, risk attitude, being unemployed and liquidity constrained are quite large, given that only about 40% of households in our sample participate in the stock market. These results allow us to conclude that these variables are the most significant determinants of the participation decision besides age and wealth.

Figure 1: Age profile of the probability to hold stocks

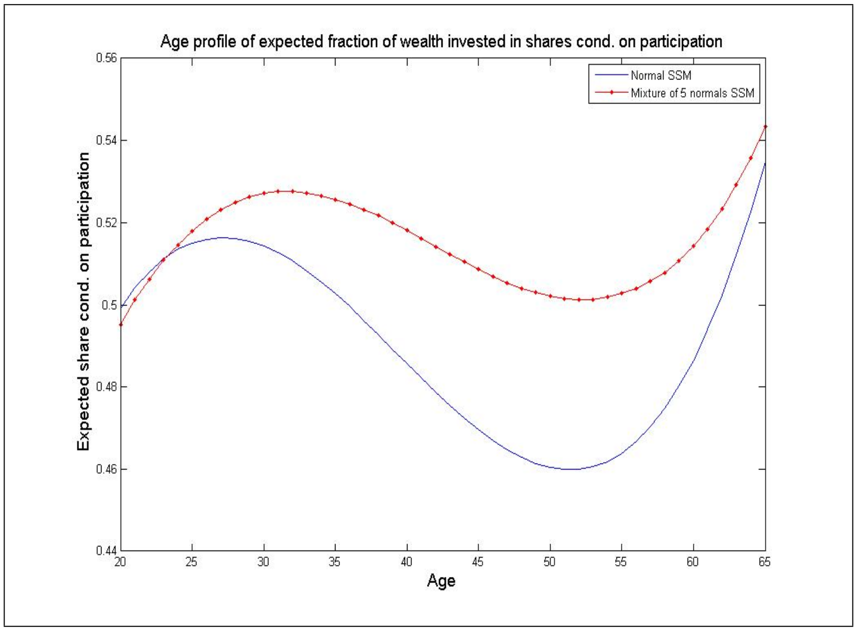


Finally, variables measuring background risk have expected effects on the stock market participation. For example, those who report large subjective uncertainty about the future of their job in the next year are almost 3% less likely to hold shares. Effect of business equity is also negative, but rather small. Increasing superannuation assets by $10,000 results in 1% increase in participation probability, which implies that households in general do not view their pension wealth as a substitute for direct investments in the 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, with most of the effects close to zero. This implies that while observed variables seem to have large explanatory power with respect to participation decision, they have modest predictive power with respect to share of wealth invested in risky assets conditional on participation. This result is similar to the findings of several studies of household portfolios in other countries (e.g. Vissing-Jorgenson, 2004). The median of the age profile for the share of wealth invested in stocks conditional on participation is shown in Figure 2. The normal model systematically underpredicts the share of risky assets at all ages and exhibits larger variability across age groups compared to the mixture model.

Increase in net worth and income have small positive effects of the share variable, but conditional on participation these effects are negative. While the negative marginal effects

Figure 2: Age profile of the fraction of wealth invested in shares conditional on participation



of net worth conditional on participation might seem surprising, it can be rationalized by the presence of unobserved heterogeneity in the data. For example, Gomes and Michaelides (2006) show in the context of the life-cycle asset allocation model that in presence of fixed cost preference heterogeneity can generate endogenous selection of more risk averse households into stock market participation. Because the degree of risk aversion determines prudence, more risk averse households will accumulate more wealth and, as a result, will be more willing to pay the fixed cost of stock market participation. Therefore, stock market participants will tend to be both wealthier and more risk averse at the same time which explains why they might invest a smaller fraction of their wealth in stocks.

In general all unconditional marginal effects of covariates on the share variable in Table 5 have expected signs. Education dummies have positive effect on the share of risky assets, increasing it by 3% to 5%, while being unemployed or out of labor force reduces share by 2% to 4%. Planning horizon and risk attitudes have strong positive effects, while subjective probability of losing a job and liquidity constraints reduce the share of risky assets in financial wealth. Comparing these results to the results of the normal model (Table 6) we observe that normal model eliminates the effect of the out of labor force dummy (posterior mean is positive but close to zero), and there substantial differences in the magnitudes of some of the marginal effects.

6 Measuring stock market participation cost

Existence of fixed cost is generally accepted as one of the main reasons for low stock market participation rates typically observed in the data. The goal of this section is to evaluate the extent of these costs in Australia using an approach similar to the one suggested by Vissing-Jorgensen (2004). In order to achieve this goal we postulate a simplified model of portfolio choice and use its predictions to construct a lower bound of the per-period costs of stock market participation.

Suppose that each household must allocate its wealth W_i between a risky asset (common stocks) with net stochastic return r_i and a riskless asset with constant return which is normalized to zero. Let α_i denote the optimal share of financial wealth invested in the risky asset and F_i denote the stock market participation cost. The participation costs measure here is intended to include both the monetary cost of participation (e.g. mutual fund fees) and the opportunity cost of time and effort devoted to understanding and processing information about the stock market (Campbell, 2006). Household i will hold a positive amount of risky asset if

$$EU[(1+\alpha_i r)(W_i - F_i)] > U(W_i) \tag{32}$$

We can also define the certainty equivalent return to the risky asset r^{ce} by the following equality

$$EU[(1+\alpha_i r)(W_i - F_i)] = U[(1+\alpha_i r^{ce})(W_i - F_i)].$$

Then we can use (32) to write the stock market participation condition as

$$(\alpha_i r^{ce} + 1)(W_i - F_i) > W_i,$$

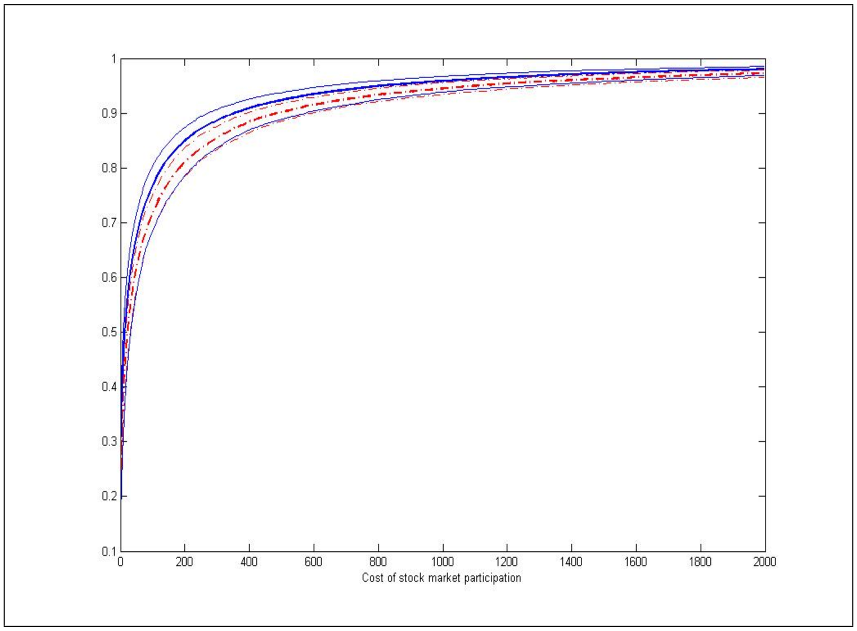
or

$$B(\alpha_i, W_i) = \frac{\alpha_i r^{ce} W_i}{1 + \alpha_i r^{ce}} > F_i. \tag{33}$$

Conversely, for non-participating households it must be the case that $B(\alpha_i, W_i) < F_i$. In other words,the benefit of stock market participation, $B(\alpha_i, W_i)$, defines the *lower bound* on the magnitude of the participation cost of a non-participating household. In order to evaluate $B(\alpha_i, W_i)$ we need to assume what value r^{ce} is likely to take. In our computations we follow Vissing-Jorgensen (2004) and set $r^{ce} = 0.04$. The sample selection model is then used to predict the optimal share of financial wealth invested in stocks, α_i, for each non-participating household.

We plot the median and the 90% confidence bounds of the empirical cdf of the expected

Figure 3: CDF of the lower bound on the stock market participation cost, Blue - Normal model, Red - Five-component mixture model.



lower bound on the cost of stock market participation below. These quantities were computed as follows:

1. Construct a grid of points from 0 to 2000;
2. For every draw from the posterior distribution of the parameters predict lower bound on the cost of stock market participation using expression (33) for individuals who have zero fraction of financial wealth invested in shares;
3. For every draw from the posterior distribution of the parameters compute proportion of individuals whose predicted cost of participation are less than the value of the grid point, for every grid point. This way for every grid point we obtain the posterior distribution of the proportion of individuals whose lower bound on the stock market participation cost is less than the value of the grid point;
4. For every point of the grid compute 10th, 50th and 90th percentiles of the posterior distribution of the proportion of individuals whose lower bound on the stock market participation cost is less than the value of the grid point. Plot the grid points against these 10th, 50th and 90th percentiles to obtain Figure 1.

Figure 3 plots the 10th, 50th and 90th of the empirical cdf of $B(\alpha_i, W_i)$ for both normal and five component mixture models. As can be seen in the figure, the normal model systematically underpredicts the magnitude of the fixed cost compared to the mixture specification. Results based on our preferred mixture model imply that participation costs needed to rationalize the fact that many household in our sample do not hold public equity need to not be very high. In particular, the cost in the region of $500-$700 will explain non-participation of 90% of the households. On the other hand, the higher percentiles of the cost distribution are quite large, which implies that factors other than the stock market participation costs might be needed to fully explain non-participation of the households in the upper right tail of the wealth distribution.

7 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 generalized sample selection 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 the generalized mixture model over the standard normal selection model, with the five-component mixture being the preferred model according to the log marginal likelihood criterion.

Results based on five-component mixture model imply that in addition to education, risk aversion, planning horizon and other household characteristics, such background risk measures as business equity and uncertainty about future labor income have significant effect on stock market participation and share of wealth invested in stocks. Compared to normal model, mixture specification produces different marginal effects of many variables as well as different age profile for participation rate. This shows that when applied to the data on household financial behavior normality assumption can be too restrictive and might lead to unreliable results. Similar to other empirical studies of household portfolios, we find that, conditional on participation, household characteristics and background risk variables have little effect on the share of wealth invested in stocks.

The paper also uses the predicted share of risky asset in total portfolio for non-participating households to construct the posterior distribution of the cost of stock market participation. We find that the normal model tends to systematically underpredict the magnitude of the fixed cost of stock market participation compared to the mixture specification. Results of the mixture model imply that a relatively modest participation cost (between \$500 and \$700) can explain non-participation of up to 90% of households in our sample.

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.

F. Gomes and A. Michaelides. Optimal life-cycle asset allocation: Understanding the empirical evidence. *Journal of Finance*, 60(2):869–904, 2006.

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

A. Nobile. Comment: Bayesian multinomial probit models with a normalization constraint. *Journal of Econometrics*, 99(2):335–345, 2000.

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.

M. van Hasselt. Bayesian inference in the sample selection and two-part models. *Working Paper*, The University of Western Ontario, 2007.

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.

Table 3: Coefficients and Marginal Effects for Participation, Mixture Model

	Mean Cf	**Std.Cf**	**Pr(Cf>0|D)**	**Mean ME**	**Std. ME**	**Pr(ME>0|D)**
const 1	-1.2466	0.9288	0.09			
const 2	0.6818	1.5438	0.69			
const 3	1.933	0.6092	1			
const 4	-1.2506	0.9672	0.10			
const 5	-0.2552	1.6567	0.41			
const	-7.3221	1.2835	0			
age	0.3249	0.0862	1	-0.0025	0.0019	0.1
age2	-0.7741	0.2067	0			
age3	0.602	0.1588	1			
nw	0.0277	0.0028	1	0.0082	0.0007	1
nw2	-0.012	0.002	0			
nw3	0.0016	0.0003	1			
income	0.0248	0.014	0.96	0.0072	0.0037	0.97
income2	-0.0006	0.0005	0.15			
edub	0.314	0.0761	1	0.11	0.0266	1
edud	0.2111	0.0634	1	0.0729	0.0217	1
eduhs	0.3122	0.0944	1	0.1094	0.0336	1
olf	-0.2051	0.0909	0.01	-0.0705	0.0305	0.01
unemployed	-0.3246	0.17	0.02	-0.1076	0.0523	0.02
liq1	-0.3831	0.0729	0	-0.1275	0.0235	0
lone	-0.0762	0.0987	0.22	-0.0271	0.0352	0.22
couple	-0.0636	0.0719	0.19	-0.0228	0.0258	0.19
children04	0.0621	0.054	0.88	0.0223	0.0194	0.88
children514	0.0048	0.0339	0.56	0.0016	0.012	0.56
prob	-0.0772	0.0564	0.08	-0.0272	0.0198	0.08
horizon1	0.2692	0.0784	1	0.0977	0.028	1
horizon2	0.2527	0.0653	1	0.0918	0.0235	1
risk	0.5597	0.0935	1	0.202	0.0314	1
bizeq	-0.0044	0.0019	0	-0.0016	0.0007	0
super	0.0289	0.0085	1	0.0086	0.0025	1
super2	-0.084	0.024	0			
super3	0.0497	0.0154	1			
nesb	-0.3264	0.081	0	-0.1098	0.0261	0
urban	0.0222	0.0556	0.65	0.0079	0.0198	0.65
st vic	-0.0827	0.0702	0.12	-0.0291	0.0247	0.12
st qld	-0.1783	0.0742	0.01	-0.0619	0.0255	0.01
st sa	0.0001	0.0934	0.50	0.0002	0.0332	0.5
st wa	-0.1633	0.0915	0.04	-0.0567	0.0314	0.04
st act	0.0842	0.2051	0.66	0.0313	0.073	0.66
st tas	-0.223	0.1456	0.06	-0.0757	0.0483	0.06

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

	Mean Cf	Std.Cf	Pr(Cf>0|D)	Mean ME	Std. ME	Pr(ME>0|D)
const	-2.6503	0.7012	0			
age	0.1101	0.0535	0.98	0.0009	0.0016	0.72
age2	-0.2706	0.1308	0.02			
age3	0.2252	0.103	0.99			
nw	0.0252	0.0022	1	0.0078	0.0006	1
nw2	-0.0114	0.0016	0			
nw3	0.0015	0.0003	1			
income	0.0277	0.0118	0.99	0.0078	0.0034	0.99
income2	-0.0008	0.0004	0.02	0	0	0
edub	0.2966	0.0685	1	0.1089	0.0254	1
edud	0.2043	0.0584	1	0.0736	0.0209	1
eduhs	0.2791	0.0843	1	0.1023	0.0316	1
olf	-0.1579	0.0795	0.02	-0.0572	0.0283	0.02
unemployed	-0.3335	0.15	0.01	-0.1143	0.0482	0.01
liq1	-0.3773	0.0678	0	-0.1299	0.0225	0
lone	-0.0735	0.0904	0.21	-0.0276	0.034	0.21
couple	-0.0638	0.0641	0.16	-0.0242	0.0243	0.16
children04	0.0812	0.0487	0.95	0.031	0.0186	0.95
children514	0.0133	0.0292	0.68	0.005	0.011	0.68
prob	-0.073	0.051	0.08	-0.027	0.0189	0.08
horizon1	0.2364	0.068	1	0.0917	0.0267	1
horizon2	0.2316	0.0568	1	0.0897	0.0222	1
risk	0.489	0.075	1	1	0.0294	1
bizeq	-0.003	0.0011	0	-0.0011	0.0004	0
super	0.0258	0.0071	1	0.0082	0.0023	1
super2	-0.0702	0.0185	0			
super3	0.04	0.0106	1			
nesb	-0.2722	0.0751	0	-0.0962	0.0255	0
urban	0.0091	0.0508	0.57	0.0034	0.0191	0.57
st vic	-0.0904	0.0626	0.07	-0.0334	0.0231	0.07
st qld	-0.1748	0.0668	0.01	-0.0634	0.0241	0.01
st sa	0.009	0.0822	.55	0.0035	0.0309	0.55
st wa	-0.1568	0.0799	0.02	-0.0569	0.0288	0.02
st act	0.0721	0.17	.66	0.0286	0.0646	0.66
st tas	-0.2026	0.1294	0.06	-0.0718	0.0447	0.06

Table 5: Coefficients and **Conditional** Marginal Effects, Share of Risky Assets, Mixture Model

	Mean Cf	**Std.Cf**	**Pr(Cf>0|D)**	**Mean ME**	**Std. ME**	**Pr(ME>0|D)**
const 1	-0.1203	2.5956	0.49			
const 2	0.0909	1.9966	0.53			
const 3	0.59	1.785	0.63			
const 4	0.0245	2.6557	0.5			
const 5	-0.588	2.4692	0.43			
const	-0.3109	1.7851	0.42			
age	0.0398	0.0228	0.97	-0.0018	0.0013	0.08
age2	-0.0991	0.0529	0.03			
age3	0.0782	0.0395	0.97			
nw	0.0006	0.0006	0.86	-0.0003	0.0005	0.24
nw2	-0.0004	0.0004	0.11			
nw3	0.0001	0.0001	0.9			
income	-0.003	0.0026	0.12	-0.0026	0.002	0.1
income2	0.0001	0.0001	0.95			
edub	0.0217	0.0174	0.89	0.0104	0.0173	0.73
edud	0.0083	0.0164	0.7	0.0005	0.0164	0.52
eduhs	0.0126	0.0221	0.73	0.0013	0.0221	0.52
olf	0.0192	0.0201	0.83	0.0268	0.0202	0.92
unemployed	0.0243	0.0409	0.72	0.0365	0.0415	0.82
liq1	0.0388	0.0219	0.96	0.0532	0.0216	0.99
lone	-0.0569	0.0266	0.02	-0.054	0.0264	0.02
couple	-0.0422	0.0158	0.01	-0.04	0.016	0.01
children04	0.0222	0.0114	0.98	0.02	0.0116	0.96
children514	-0.0021	0.0077	0.4	-0.0023	0.0077	0.38
prob	-0.0073	0.0121	0.26	-0.0046	0.0121	0.34
horizon1	-0.0195	0.0151	0.1	-0.0289	0.0151	0.02
horizon2	0.008	0.0131	0.72	-0.0007	0.013	0.48
risk	0.0704	0.0149	1	0.0521	0.0155	1
bizeq	-0.0004	0.0002	0.02	-0.0002	0.0002	0.16
super	0.0006	0.0015	0.67	-0.0002	0.0013	0.45
super2	0.0006	0.0037	0.57			
super3	-0.0006	0.0021	0.38			

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

	Mean Cf	Std.Cf	Pr(Cf>0|D)	Mean ME	Std. ME	Pr(ME>0|D)
const	-0.1387	0.433	0.36			
age	0.0429	0.0294	0.93	-0.0032	0.0017	0.03
age2	-0.1174	0.0684	0.05			
age3	0.0993	0.0513	0.97			
nw	0.0015	0.0013	0.89	-0.0006	0.0006	0.13
nw2	-0.0009	0.0007	0.1			
nw3	0.0001	0.0001	0.91			
income	-0.0018	0.0038	0.32	-0.0031	0.0028	0.13
income2	0.0001	0.0001	0.72			
edub	0.0654	0.0269	0.99	0.0401	0.0239	0.95
edud	0.0219	0.0241	0.82	0.0041	0.0221	0.58
eduhs	0.0691	0.0332	0.98	0.0451	0.0304	0.93
olf	0.0754	0.0295	0.99	0.0887	0.0284	1
unemployed	0.0125	0.0681	0.57	0.0421	0.0648	0.75
liq1	0.0662	0.041	0.96	0.0996	0.0306	1
lone	-0.0557	0.0362	0.06	-0.0492	0.0352	0.08
couple	-0.064	0.0231	0	-0.0583	0.0221	0
children04	0.0214	0.0172	0.9	0.0145	0.0166	0.81
children514	0.0193	0.0103	0.97	0.0181	0.0099	0.96
prob	-0.009	0.018	0.32	-0.0027	0.0176	0.44
horizon1	-0.0324	0.0245	0.1	-0.052	0.0229	0.02
horizon2	0.0264	0.0204	0.9	0.0073	0.0183	0.66
risk	0.1408	0.0266	1	0.102	0.0226	1
bizeq	-0.0006	0.0003	0.02	-0.0003	0.0003	0.13
super	0.0047	0.0024	0.97	0.0025	0.0018	0.91
super2	-0.0071	0.0059	0.11			
super3	0.0028	0.0032	0.82			

Table 7: Coefficients and **Unconditional** Marginal Effects, Share of Risky Assets, Mixture Model

	Mean Cf	**Std.Cf**	**Pr(Cf>0|D)**	**Mean ME**	**Std. ME**	**Pr(ME>0|D)**
const 1	-0.1203	2.5956	0.49			
const 2	0.0909	1.9966	0.53			
const 3	0.59	1.785	0.63			
const 4	0.0245	2.6557	0.5			
const 5	-0.588	2.4692	0.43			
const	-0.3109	1.7851	0.42			
age	0.0398	0.0228	0.97	-0.0019	0.0011	0.03
age2	-0.0991	0.0529	0.03			
age3	0.0782	0.0395	0.97			
nw	0.0006	0.0006	0.86	0.004	0.0004	1
nw2	-0.0004	0.0004	0.11			
nw3	0.0001	0.0001	0.9			
income	-0.003	0.0026	0.12	0.0026	0.0019	0.9
income2	0.0001	0.0001	0.95			
edub	0.0217	0.0174	0.89	0.0598	0.0146	1
edud	0.0083	0.0164	0.7	0.0369	0.0118	1
eduhs	0.0126	0.0221	0.73	0.0559	0.0185	1
olf	0.0192	0.0201	0.83	-0.0272	0.0165	0.05
unemployed	0.0243	0.0409	0.72	-0.0449	0.029	0.07
liq1	0.0388	0.0219	0.96	-0.0512	0.013	0
lone	-0.0569	0.0266	0.02	-0.0351	0.0196	0.03
couple	-0.0422	0.0158	0.01	-0.0272	0.014	0.03
children04	0.0222	0.0114	0.98	0.0193	0.0103	0.97
children514	-0.0021	0.0077	0.4	-0.0001	0.0063	0.51
prob	-0.0073	0.0121	0.26	-0.015	0.0106	0.07
horizon1	-0.0195	0.0151	0.1	0.0354	0.0142	1
horizon2	0.008	0.0131	0.72	0.0457	0.0127	1
risk	0.0704	0.0149	1	0.1324	0.0185	1
bizeq	-0.0004	0.0002	0.02	-0.0009	0.0003	0
super	0.0006	0.0015	0.67	0.0042	0.0013	1
super2	0.0006	0.0037	0.57			
super3	-0.0006	0.0021	0.38			

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

	Mean Cf	**Std.Cf**	**Pr(Cf>0|D)**	**Mean ME**	**Std. ME**	**Pr(ME>0|D)**
const	-0.1387	0.433	0.36			
age	0.0429	0.0294	0.93	-0.0008	0.0009	0.18
age2	-0.1174	0.0684	0.05			
age3	0.0993	0.0513	0.97			
nw	0.0015	0.0013	0.89	0.0032	0.0004	1
nw2	-0.0009	0.0007	0.1			
nw3	0.0001	0.0001	0.91			
income	-0.0018	0.0038	0.32	0.0023	0.0017	0.92
income2	0.0001	0.0001	0.72			
edub	0.0654	0.0269	0.99	0.0637	0.0144	1
edud	0.0219	0.0241	0.82	0.0334	0.0115	1
eduhs	0.0691	0.0332	0.98	0.0626	0.0189	1
olf	0.0754	0.0295	0.99	0.003	0.0166	0.57
unemployed	0.0125	0.0681	0.57	-0.0394	0.029	0.1
liq1	0.0662	0.041	0.96	-0.0334	0.0137	0.01
lone	-0.0557	0.0362	0.06	-0.0317	0.02	0.06
couple	-0.064	0.0231	0	-0.0332	0.0144	0.01
children04	0.0214	0.0172	0.9	0.0193	0.0102	0.97
children514	0.0193	0.0103	0.97	0.0091	0.0059	0.93
prob	-0.009	0.018	0.32	-0.0129	0.0099	0.09
horizon1	-0.0324	0.0245	0.1	0.0165	0.0144	0.88
horizon2	0.0264	0.0204	0.9	0.0427	0.0123	1
risk	0.1408	0.0266	1	0.1417	0.0204	1
bizeq	-0.0006	0.0003	0.02	-0.0006	0.0002	0
super	0.0047	0.0024	0.97	0.0045	0.0012	1
super2	-0.0071	0.0059	0.11			
super3	0.0028	0.0032	0.82			