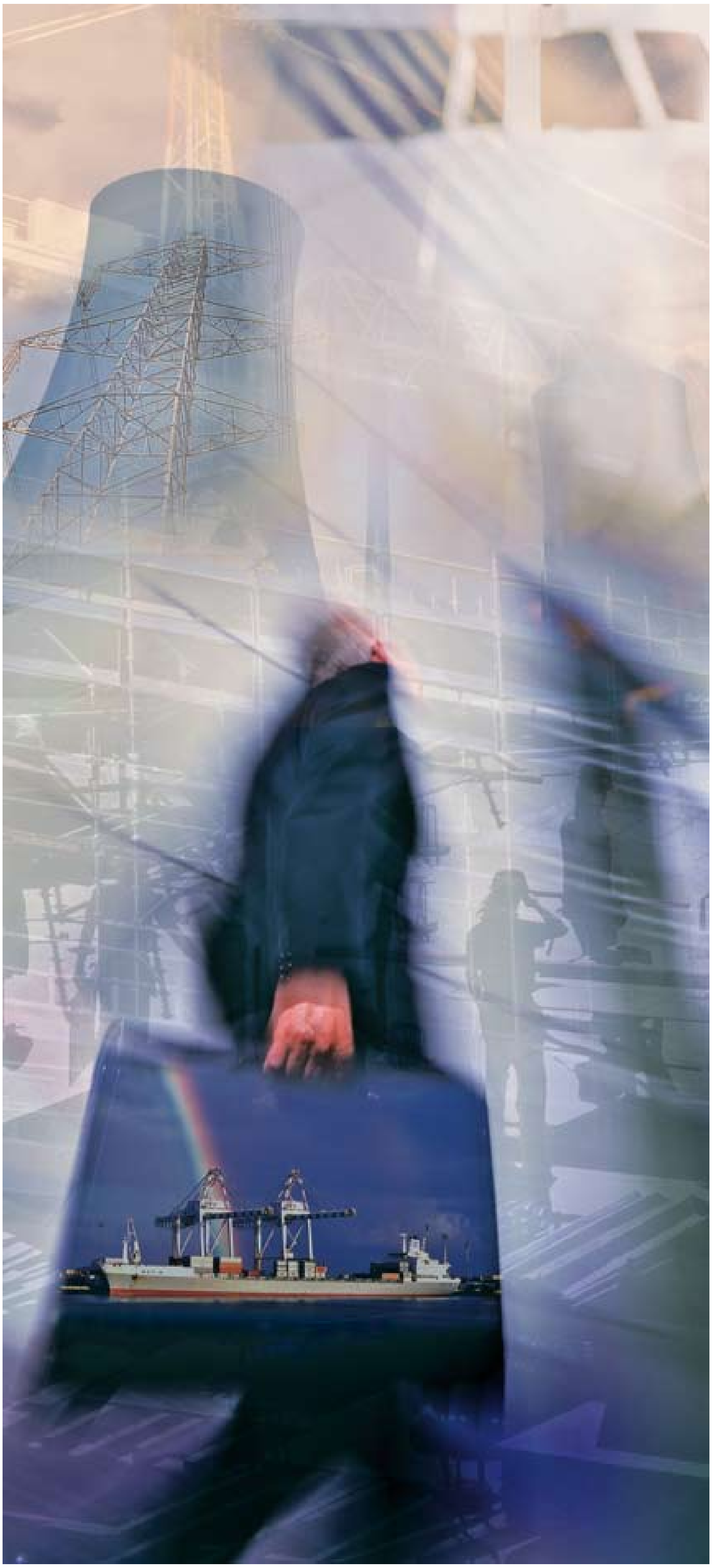


Queensland

The Economic Society of Australia Inc.

Proceedings of the 37th Australian Conference of Economists

Papers delivered at ACE 08

30th September to 4th October 2008
Gold Coast Queensland Australia

ISBN 978-0-9591806-4-0

Welcome

The Economic Society of Australia warmly welcomes you to the Gold Coast, Queensland, Australia for the 37th Australian Conference of Economists.

The Society was formed 83 years ago in 1925. At the time, the Society was opposed to declarations of policy and instead focused on open discussions and encouraging economic debate. Nothing has changed today, with the Society and the conference being at the forefront of encouraging debate.

This year we have a large number of papers dealing with Infrastructure, Central Banking and Trade.

Matters of the greatest global importance invariably boil down to be economic problems. Recent times have seen an explosion of infrastructure spending, after world-wide population growth has seen demand outpace aging supply. The world has become more globalised than at any time since World War I but the benefits of this (and the impact on our climate) has been questioned by some.

At the time of preparing for this conference we could not have known that it would have been held during the largest credit crisis since the Great Depression. The general public and politicians both look to central banks for the answers.

We are also very pleased to see a wide selection of papers ranging from applied economics to welfare economics. An A – Z of economics (well, almost).

Another feature of this conference is that we have gone out of our way to bring together economists from all walks of life, in particular from academia, government and the private sector. We are grateful to all of our sponsors, who are as diverse as the speakers.

The Organising Committee

James Dick
Khorshed Alam (Programme Chair)
Michael Knox
Greg Hall
Allan Layton
Rimu Nelson
Gudrun Meyer-Boehm
Jay Bandaralage
Paula Knight

Published November 2008

GPO Box 1170
Brisbane Queensland Australia
ecosocqld@optushome.com.au

Our Gold Sponsors

QIC
Queensland Treasury Corporation
Reserve Bank of Australia
Queensland Government Department of Infrastructure and Planning

Keynote Sponsors

Australian Government Productivity Commission
Australian Government The Treasury
Queensland University of Technology Faculty of Business School of Economics and Finance
The University of Queensland Australia
Goldman Sachs JBWere
The University of Queensland Australia
NCER National Centre for Econometric Research
ES Central Council

Special Session Sponsors

Griffith University
The Higher Education Academy Economics Network

Unless we have specifically been requested to do otherwise, all the papers presented at the conference are published in the proceedings in full. A small number of papers will have versions that have also been made available for special editions of Journals, Economic Analysis and Policy, and the Economic Record. Authors will retain the right to seek additional publication for papers presented at the conference so long as it differs in some meaningful way from those published here.

The opinions expressed in the papers included in the proceedings are those of the author(s) and no responsibility can be accepted by the Economic Society of Australia Inc, Economic Society of Australia (Queensland) Inc, the publisher for any damages resulting from usage or dissemination of this work.

The Paper following forms part of - *Proceedings of the 37th Australian Conference of Economists*
ISBN 978-0-9591806-4-0

Portfolio Allocation and the Endogeneity of Wealth in Australian Households

Elisabeth Huynh*

30th June 2008

Abstract

This paper studies the main determinants for household share of risky assets held in wealth using the Household Income Labour Dynamics Australia survey. The fractional model is proposed for the first time to portfolio allocation literature to account for the proportional dependent variable. In most cases considered the fractional model specification is favoured over the widely used tobit suggesting it is suitable for modelling portfolio holdings. Although wealth is important in explaining portfolio holdings and attitudes to risk, little attention is devoted to correcting for potential endogeneity of wealth. This study allows for endogenous wealth using a two-stage method in determining the average partial effects. Endogeneity is supported and the bias is large enough to be considered important.

*Author: PhD Student, University of Sydney, Economic and Business building (H69), Darlington, NSW 2006, Australia. Fax +61-2-9351-6409, E-mail ehuy7515@mail.usyd.edu.au

I Introduction

One of the most important decisions households will ever make is how they choose to invest their money. In 2000, world household wealth was approximately 125.3 trillion US dollars (Davies et al., 2008), and composed of large investments in housing and financial assets. In developing countries the average household portfolio consisted predominantly of investments in real assets though in the developed world there is greater emphasis on investments in financial assets. This variation in composition at the international level can largely be explained by country specific factors[1]. In the domestic setting robust patterns of investment suggest that behavioural or attitudinal distinctions may exist among households on demographic factors such as wealth, socio-economic status, ethnicity, life-cycle stage, and education. Stimulated by the need to better understand the underlying attitudes that influence investment decisions, the study of household portfolio holdings continues to be an important area in economics and finance.

Theoretical studies of household portfolios over the past forty years have contributed significantly to our understanding of how risk aversion influences the propensity of households to take financial risk. Standard portfolio theory in particular shows that household choices between risk-free and risky financial assets are directly related to risk aversion. Empirical studies estimate the share of risky assets held in wealth in an attempt to determine the levels and slopes for risk aversion (Bellante and Green, 2004; Faig and Shum, 2005). A question asked in these studies concerns whether relative risk aversion increases, decreases or remains constant with rises in wealth. The evidence is less than conclusive and is a subject of ongoing investigation. Using the Household Income Dynamics Australia (HILDA) survey this paper provides evidence on the role of heterogeneity such as gender, marital status, ethnicity, geographical and educational differences on household investments to explain risk attitudes of households.

Despite wealth being an important variable in the analysis of portfolio holdings, little attention is been paid to addressing the potential endogeneity of wealth in the literature. A partial listing of papers that treat wealth as exogenous include Bellante and Green (2004), Faig and Shum (2005) and Riley and Chow (1992). The potential endogeneity of wealth if not corrected for can bias the effect of wealth on the share of risky assets held, and consequently the form of risk aversion.

This study provides two methodological contributions to the literature. First, potential

[1] This includes differences in government policies, politics or trading emphasis across countries. For instance, Davies et al. (2008) notes that there is a greater reliance on agricultural production in developing countries, and this can result in larger investments into real estate in the form of farmland.

endogeneity of wealth is accounted for using a two-stage method outlined by Wooldridge (2005). The average partial effects are calculated and compared to those in a model with exogenous wealth to determine the endogeneity bias. Second, we extend the econometric application to a fractional model proposed by Papke and Wooldridge (1996). The fractional model accounts for a proportional dependent variable with many observations at the extreme values of zero and one. The fractional model is also arguably more appropriate to the tobit because the values outside the [0,1] interval are unobserved because they are not defined, but rather than from censoring as implicitly assumed in the tobit model. Furthermore, using quasilikelihood estimation methods the fractional model maintains desirable asymptotic properties for the estimators without the need for strong parametric assumptions. For the above reasons, this paper adopts the fractional model to model the share of risky assets and compare its performance against the double-censored tobit.

This study also contributes to the current empirical literature using newly available information on wealth provided by the Household Income and Labour Dynamics in Australia (HILDA, Wave2) survey. The HILDA survey is the first known Australian Survey since 1950s that provides extensive information on asset holdings, demographic and socioeconomic variables at the household and individual level. One advantage of HILDA is that unlike most datasets used in the literature, it allows for identification of household head.

This paper is structured as follows. In Section II we examine the literature provided by the economic theory on how portfolio shares vary with household characteristics and examine the past empirical evidence on observed behaviour of households. Section III describes the economic model, and presents the empirical modelling and identification issues. Section IV describes the empirical data in detail, and Section V presents the results of our analysis of the data. Sections VI discusses the main findings, and concludes in Section VII.

II Literature Review

Knowing the relationship between risk aversion and wealth is important for choosing the correct utility representation for consumer preferences since utility functions implicitly impose the form of risk aversion[2]. Two measures of risk aversion are proposed in the literature: the

[2]For instance, the isoelastic utility function and the power utility function exhibit decreasing absolute risk aversion and constant relative risk aversion; the quadratic utility function is increasing absolute risk aversion and increasing relative risk aversion; and the exponential utility function is constant absolute risk aversion and increasing relative risk aversion. The utility functions discussed previously belong to Hyperbolic absolute risk aversion (HARA) class of utility functions. HARA is often assumed in economic theory because it allows for the possibility of decreasing, increasing and constant relative risk aversion, but to solve a utility maximisation problem a specific utility function is often required.

Arrow-Pratt coefficient of absolute risk aversion (ARA) and relative risk aversion (RRA). Although the consensus in the literature is that ARA declines with rises in wealth, the form of RRA is less than conclusive. Arrow (1974) originally conjectured that relative risk aversion is either increasing or constant, but agrees that the intuition is not clear. Friend (1973) on the other hand suggests that decreasing relative risk aversion is also a plausible assumption. Because the form of RRA is not intuitively clear, the relationship between relative risk aversion and wealth has thus become an empirical question.

Empirical investigations on describing the relative risk aversion with changes in wealth are extensive, but conflicting. Although many studies find supporting evidence for decreasing relative risk aversion (DRRA)[3], there is also considerable support for increasing relative risk aversion (IRRA) and/or constant relative risk aversion (CRRA) (see, Morin and Suarez, 1983; Riley and Chow, 1992). Moreover, the form of RRA is sensitive to different wealth levels (see Morin and Suarez, 1983)[4] and measures for wealth and risky assets. Measuring wealth to include housing as in Bellante and Green (2004) and human capital as in Riley and Chow (1992) can affect conclusions on whether risk aversion is increasing, decreasing or constant with wealth.

The role of heterogeneity on relative risk aversion (RRA) also provides some explanation for differences in utility preferences among consumers. Jianakoplos and Bernasek (1998) find women show greater risk aversion than men do, and being employed or retired also tends to increase RRA. The evidence is less obvious for studies on the effect of age. Studies by Palsson (1996) and Bellante and Green (2004) find evidence supporting IRRA with age. Riley and Chow (1992) on the otherhand, find evidence for DRRA with age up to 65 years of age, and IRRA when they define housing wealth as risky. Other variables considered include the effect of ethnicity (Jianakoplos and Bernasek, 1998), household size, living regions (Guiso and Paiella, 2001), health (Bellante and Green, 2004), and education (Hochguertel et al., 1997; Bellante and Green, 2004). Findings in the literature suggest that heterogeneity accounts for some of the differences in the portfolio holdings.

Portfolio holdings continue to be a thriving area of research with recent extensions for taxation by Palsson (1996) and King and Leape (1998); liquidity constraints as in Guiso, Jappeli, and Terlizzese (1996), Arrondel (2000) and Guiso and Paiella (2001); housing ownership as in McCarthy (2004); and human capital in Bellante and Green (2004). To investigate all these are beyond the scope of this paper, but we address the effect of taxes on portfolio

[3] Decreasing relative risk aversion is supported in Hochguertel et al. (1997); Jianakoplos and Bernasek (1998); Faig and Shum (2005).

[4] Morin and Suarez (1983) in their study on Canadian households find IRRA for low wealth households, DRRA for households with higher wealth below \$100,000, and CRRA for wealth above \$100,000.

holdings. The idea is that the differences in effective tax rates among assets may lead to portfolio specialisation because it may alter the tradeoff between risk and return.

Finally, although wealth is important in explaining portfolio allocation, few have extended the analysis for potential endogeneity of wealth in the literature. There appear to be two studies that account for the endogeneity of wealth. The first study is Hochguertel et al. (1997) who implicitly account for endogeneity of wealth on the decision to hold assets. Their study however does not compare findings to a model with exogenous wealth or report the endogeneity bias. Second, Guiso and Paiella (2001) consider endogenous wealth and consumption in their study on the effect of wealth and background risk on a direct measure of absolute risk aversion. Their results find that the endogeneity bias of wealth is towards zero. The existing research has suggested that wealth is endogenous in portfolio holdings. Thus, this study accounts for the endogeneity of wealth and determine the endogeneity bias.

III Model

This paper adopts the classically formalised optimal portfolio choice problem of investor expected utility maximisation. The optimal demand for risky assets in wealth α is given by[5],

$$\alpha \backsimeq (1/\gamma)\, MPR, \tag{1}$$

where MPR is the market value of all risky assets, also known as the market price of risk, and γ is the Arrow-Pratt measure of relative risk aversion. This result provides a simple expression for describing the share of risky asset to wealth, α as a function of the MPR and the inverse of the Arrow-Pratt coefficient of relative risk aversion γ. The empirical application of equation (1) simply involves measuring α and estimating it against wealth and variables that control for heterogeneity. Under the inclusion of taxes, the right hand side of equation (1) is replaced with the tax adjusted demand for risky asset, given by α multiplied by one minus taxes.

To quantify risk aversion in households, we posit the following relation for the share of risky assets to wealth:

$$\alpha = G\left(W, \mathbf{z_1}\right) + \epsilon,$$

where W is current wealth, $\mathbf{z_1}$ is a vector of variables which influences the age-wealth profile and ϵ is the error term. The dependent variable is the dollar value of risky assets per dollar of

[5]For the derivation of the equations, see for example Siegel and Hoban (1982).

wealth. The vector $\mathbf{z_1}$ includes socio-economic variables for tastes and family structure (age, education, country of birth, marital status, gender, employment) and economic environment (geographical location) of the household. Similar specifications have been used by Riley and Chow (1992), Guiso, Jappeli, and Terlizzese (1996) and Bellante and Green (2004).

General observations from the literature invite several hypothesis relating to the share of risky assets to wealth and socio-economic variables. First, we predict a non-linear relationship between W and the share of risky assets. In particular, we model W as a quadratic function following Bellante and Green (2004) and Guiso, Jappeli, and Terlizzese (1996).

Second, wealth is hypothesised to be an endogenous variable. The direction of the endogeneity bias, is not clear a priori. If more risk averse individuals choose safer but less rewarding prospects, they may end up poorer. This would overstate the positive relation between the proportion of risky assets held and wealth. However, if more risk averse individuals are also more prudent, ceteris paribus, they will compress current consumption, save more and end up accumulating more assets. In this case, our estimates of the relationship between the share of risky assets and wealth will be biased towards zero.

With socio-economic variables, we further hypothesise the following: men hold a greater proportion of risky asset to wealth thereby, showing greater risk tolerance than women. Married households more likely to hold less risky assets to smooth their risk. To the extent that greater education can reflect greater awareness and sophistication of risk and return benefits of holding risky types of assets, we conjecture that the share of risky assets increases with education.

Estimation

We consider four different measures for the dependent variable that closely reflect the empirical work in this area, and this is outlined in detail in section IV. For each of these measures, we estimate the share of risky asset to wealth using a double-censored tobit procedure and compare the results to those from the fractional logit model (FLM) and fractional probit model (FPM) estimations. To determine the size of the effects, we report the mean average partial effects for the variables considered. In addition, we re-estimate the double-censored tobit and FPM allowing for the potential endogeneity of wealth and report the APE.

Fractional logit and probit model. The share of risky asset held in wealth α is naturally censored at and including zero and one. An observed proportion equal to zero represents an individual's choice not to invest in risky assets, but allocate all their wealth to other types of assets such as non-risky or real assets. An observed proportion equal to one reflects the choice to invest all wealth in risky assets. In wealth holding data, the large portion of zero holdings of risky assets is often observed despite evidence of an 'equity premium' in the market, pertaining to the "equity holding puzzle".

The double-censored tobit is a natural model for the current problem. However, one criticism when modelling the tobit to proportional data are that values outside the [0, 1] range are not caused by censoring, but because they are not defined outside this interval. Further, the proportions are not normally distributed because they are not defined over the real space (see Kieschnick and McCullough, 2003, for a further discussion).

The fractional response model as outlined by Papke and Wooldridge (1996) is appropriate when the dependent variable α is defined in the interval [0, 1], and can take many observations at the extreme values of zero and one. Under this model specification, the share of risky asset to wealth for the ith household is given by

$$E\left[\alpha_i \mid \mathbf{x}_i\right] = G\left(\mathbf{x}_i\beta\right) \;\; i = 1, 2, ..., n, \tag{2}$$

where $G\left(.\right)$ is a known function satisfying $0 < G\left(z\right) < 1$ for all $z \in \mathbb{R}$, and $\mathbf{x_i}$ is a 1xk vector of observables. For the fractional logit model, $G\left(.\right)$ is assumed to be a logistic function, and follows the normal cdf in the case of the fractional probit model. Papke and Wooldridge (1996) propose to estimate the parameters using a Bernoulli quasi-likelihood method as in Gourieroux, Monfort, and Trognon (1984). The quasi-likelihood procedure is arguably more flexible to the maximum likelihood procedure in that it does not require a full parametric model for the dependent variable, but only the first and second moments, and the estimates possess good asymptotic properties. Provided that the mean specification (2) holds, the maximum quasi-likelihood estimate of β is always consistent. If the variance is also correctly specified, the estimate of β is asymptotically normal, unbiased and efficient. When the mean is correctly specified but the variance function is not, the parameter estimates are consistent and asymptotically unbiased and normal but inefficient. To account for this potential model misspecification, a robust estimator for the asymptotic variance as in Papke and Wooldridge (1996) will be used in this study.[6]

[6] The robust Papke and Wooldridge (1996) variance estimator is given by,

$$\left(\mathbf{D'V^{-1}D}\right)^{-1}\left(\mathbf{DV^{-1}EV^{-1}D}\right)\left(\mathbf{D'V^{-1}D}\right)^{-1}, \tag{3}$$

where $\mathbf{V}$ is the specified variance and $\mathbf{D}$ is an nxk matrix containing the partial derivative of (2) with respect to β, and the matrix $\mathbf{E} = diag\left[\left(\alpha_1 - G\left(\mathbf{x_1}\beta\right)\right)^2, \left(\alpha_2 - G\left(\mathbf{x_2}\beta\right)\right)^2, \ldots, \left(\alpha_n - G\left(\mathbf{x_n}\beta\right)\right)^2\right]$ evaluated at the quasi-likelihood estimate $\hat{\beta}$.

Average partial effects. Suppose we explicitly allow the expectation of the response variable α, to depend on unobserved heterogeneity. We consider stating model of interest as $E(\alpha \mid \mathbf{x}, \theta) = G(x_i\beta + \theta)$, where θ is unobserved heterogeneity independent of $\mathbf{x}$[7]. Various factors including financial ability might be contained in θ. When the expectation of the response variable α, explicitly depends on the unobserved heterogeneity, such that $E(\alpha \mid \mathbf{x}, \theta) \neq E(\alpha \mid \mathbf{x})$, then the changes of $\mathbf{x}$ on α cannot be characterised by the marginal effects of $\mathbf{x}$ on equation (2). Wooldridge (2005) suggests studying the changes in $\mathbf{x}$ on α holding the unobserved heterogeneity fixed. This leads to estimating the average partial effects (APE). In the case of an endogenous explanatory variable in $\mathbf{x}$, the unobserved heterogeneity is correlated with the endogenous explanatory variable and the APE is required to untangle the effects of θ on α from $\mathbf{x}$ on α.

Wooldridge (2005) outlines a two step procedure for estimating APE to account for possible endogenous explanatory variables for the random coefficient model, probit and tobit model. To date, this is the first study to estimate the APEs for fractional response models, and moreover assuming potential endogenous explanatory variables. Under the normality assumptions and properties, the APE for the case of the fractional probit model can be derived.[8] As in Wooldridge (2005) we assume there exists an h-vector of observable covariates $\mathbf{w}$ such that θ and $\mathbf{x}$ are independent on $\mathbf{w}$, and redundant in the structural expectation such that the model can be written as

$$E(\alpha \mid \mathbf{x}, \mathbf{w}, \theta) = \Phi(\mathbf{x}\beta + \theta),$$

where $\mathbf{x} \equiv (\mathbf{z_1}, y_2)$ and $\mathbf{z}_1$ is the m-vector of observed exogenous covariates for α uncorrelated with $\mathbf{w}$, and y_2 is the observed endogenous explanatory variable. The term θ is independent of $\mathbf{z_1}$, and contains unobserved heterogeneity assumed to be correlated with all elements of y_2.

The following relationship is assumed for the endogenous explanatory variable,

$$y_2 = \mathbf{z}\delta + \mathbf{w} \quad \mathbf{w} \mid \mathbf{z} \sim N\left(0, \sigma^2_{\mathbf{w}}\right), \tag{4}$$

where y_2 is a well-defined function of $(\mathbf{z}, \mathbf{w})$ and $\mathbf{z}$ is a vector of observed exogenous covariates including covariates in $\mathbf{z_1}$ and exogenous variables other than those included in $\mathbf{z_1}$ correlated

[7] All equations are represented at the household level, we have removed the household ith subscript.

[8] This is not the case for the FLM, although arguably the fractional probit specification can be considered to be a close approximation of the logistic one.

with y_2 and independent of θ. The variance of θ is σ_1^2, and $\sigma_{\mathbf{w}}^2$ for $\mathbf{w}$. Joint normality of $(\theta, \mathbf{w})$ is assumed with a covariance of $E(\mathbf{w}'\theta)$.

With the independence from $\mathbf{z}$, we can express the unobserved heterogeneity in terms of $\mathbf{w}$,

$$\theta = \mathbf{w}\gamma + e,$$

where $\gamma = \sigma_{\mathbf{w}}^{-1} E(\mathbf{w}'\theta)$ is a h-vector of parameters, such that e is independent of $(\mathbf{x}, \mathbf{w})$ with a normal $(0, \sigma_e^2)$ distribution, where $\sigma_e^2 = \sigma_1^2 - \sigma_{\mathbf{w}}^2 \gamma'\gamma$.

Under the above assumptions and correcting for neglected heterogeneity, a consistent estimator for the APE of x_j evaluated at $\mathbf{x^0}$ is,

$$\hat{\beta}_j^* \sum_{i=1}^{N} \phi\left(\mathbf{x^0}\hat{\beta}^* + \mathbf{w}_i\hat{\gamma}^*\right)/N, \tag{5}$$

where $\hat{\beta}^*$and $\hat{\gamma}^*$ are the consistent estimates from a simple fractional probit analysis of α on $\mathbf{x}$ and $\mathbf{w}$. The presence of $\mathbf{w}$ controls for endogenous variation in $\mathbf{x}$, such that the omitted heterogeneity is uncorrelated with $\mathbf{x}$ in a probit analysis of α on $\mathbf{x}$ and $\mathbf{w}$.

Since $\mathbf{w}$ is unobserved, a two step approach for estimating APE is used. The first stage involves estimating the equation (4) by ordinary least squares (OLS) and saving the residuals $\hat{w}_i$ for $i = 1, 2, ..., N$. The second stage involves estimating the fractional probit model of α on $\mathbf{x}$ and $\hat{w}_i$, $i = 1, 2, ..., N$. The coefficient estimates and the robust standard errors are calculated and used to estimate the APE in equation (5).

IV Data and data description

The present study uses household wealth information from Wave 2 of the Household Income and Labour Dynamics of Australia (HILDA)[9]. It comprises of 13,041 persons from 7245 households. HILDA is a household-based panel survey that follows all members of an initial sample annually over an indefinite time. Each wave in HILDA provides general information on the household and its member, including a range of personal characteristics such as age, gender, work histories, labour market status, income and marital status. Moreover, Wave 2 of the HILDA survey conducted in 2002 also poses questions related to wealth, so providing what is believed to be the first-major household-level survey of wealth in Australia.

[9]HILDA is initiated and funded by the Australian Government through the Department of Families, Community Services and Indigenous Affairs for analysis.

Like most studies, this survey has its limitations. First, this survey appears to suffer from underreporting of household asset amounts because asset values are reported by the individual and therefore may not be accurate. Second, although the survey contains extensive information on assets and liabilities amounts, the survey does not report on ownership of single asset units. The data provide subtotal quantities for categories of assets including real and financial assets, liabilities and mortgages. We aggregate assets in categories that are somewhat homogeneous with respect to their risk profile and over all respondents per household.

To appropriately study heterogeneity in risk attitudes, this study attempts to correctly identify the household head. The HILDA survey questions respondents on the household head in financial decision-making. Cases in which responses appeared clear and consistent among all household members have only been considered in this analysis. This reduced the sample size to 3660 cases[10].

This paper follows a standard definition for risky assets consistent with the main body of empirical literature in this area. A risky asset is defined as an asset that provides an uncertain nominal cash flow and therefore includes risky financial and real assets. Risky financial assets follow the simplest definition of direct stockholding and some risky assets held indirectly through mutual funds and retirement accounts. Risky real assets include risky business and property related assets. Total risky assets are measured as the sum of financial equity, business equity, and equity in property other than own home[11].

Recently there has been a wide concern on the impact of home ownership on household portfolio allocation of wealth. A substantial part of wealth of homeowners is held in housing wealth. To incorporate for this, we also consider including housing wealth as a risky asset. As in Bodie, Merton, and Samuelson (1992) housing wealth is measured as the equity of one's own home.

— insert Table 1 about here —

We therefore analyse four dependent variables related to the portfolio share of risky assets in wealth. (1) *Alpha1:* share of risky asset with housing included as a risky asset and (2) *Alpha2:* Tax-adjusted share of risky asset under the inclusion of housing as a risky asset,

[10]Refer to Appendix 1 for a detailed outline of how the household head was identified from the data.

[11]Financial equity is defined as risky financial assets comprising of equity investments including shares, managed funds and, property trusts, net of liabilities and excluding mortgage debt. Business equity is defined as business value less any business debt owed by the household on the business, where business value includes the gross value of any property, buildings, vehicles, machinery and bank accounts owned by the business.

(3) *Alpha3:* share of risky asset excluding housing as a risky asset, and (4) *Alpha4:* share of risky asset excluding housing and adjusted for tax. Wealth is measured as net worth.

Table 1 presents the descriptive statistics and explanation of the key variables for the full sample. The share of risky assets held in wealth is larger for *Alpha1* and *Alpha2*, averaging 52% and 44% of the sample respectively. For *Alpha3* and *Alpha4*, the risky asset holding is low averaging 16% and 13% respectively. From Table 1, 73% of household heads graduated from high school, and 61% lived in a major city as opposed to inner and outer regional, remote and migratory. There is a fairly equal proportion of males and females in the sample. Age of the household head averages around 46 years of age and retired household heads comprise of 19% of the sample. Finally, household wealth averages around \$390,000 and maximum wealth (top-coded) at \$3 million dollars.

—insert Table 2 about here—

Table 2 presents the proportion of wealth held in risky assets over four quantiles of wealth. First, we observe a large portion of low wealth households, with three quantiles holding wealth below \$500,600. The measure of wealth in this context is net worth and includes debt and liabilities, and that to some extent may explain for low wealth levels in the data. Second, the proportion of wealth held in risky assets rises with the wealth of the investor and consistent with decreasing relative risk aversion (DRRA). For *Alpha2* and *Alpha4* where housing is excluded, we further observe DRRA up to the 3rd quantile of wealth of \$509,600 and increasing relative risk aversion (IRRA) thereafter. This suggests that the relationship between wealth and the share of risky assets is potentially quadratic and concave.

The HILDA survey also provides a subjective measure for financial risk taking, which was used in this study to construct a simple measure for financial risk aversion. Table 2 shows how the proportion of financially risk averse individual's decreases over wealth categories. This provides further support for DRRA.

V Results

Table 3 presents the estimation results with exogenous wealth for the fractional logit (FLM), fractional probit (FPM) and the double-censored tobit model (Tobit), and separated into the four different measurements considered for the dependent variable.

— Insert Table 3 about here —

The estimated coefficient on *Wealth* provides an estimate of the inverse of the coefficient of relative risk aversion up to a positive multiplicative constant. A positive (negative) coefficient indicates decreasing (increasing) relative risk aversion. The variables of interest *Wealth* and *Wealth2* are highly significant in all the models and definitions considered. The coefficient for *Wealth* is positive, indicating DRRA, and *Wealth2* is negative, indicating a concave quadratic relationship for wealth. Holding other variables constant, households have DRRA up to 1.64-1.68 millions of dollars of wealth and IRRA for greater holdings of wealth for Models (1)-(3) with housing. When housing is excluded as in Models (7)-(12), households show DRRA up to 2.3-2.8 millions of dollars of wealth and IRRA for greater holdings of wealth. The adjustment of taxes in Models (4)-(6) and (10)-(12), dampens the effect of wealth and households show DRRA up to 1.53-1.54 and 2.2-2.6 millions of dollars of wealth including and excluding housing, respectively. Figure 1 presents the graphs for the predicted share of risky assets on wealth for the tobit, fractional logit and fractional probit specifications for *Alpha1-4*. Figure 1 shows that the relationship between wealth and the the share of risky assets is concave. Although the shapes across model specifications are similar, there are observable differences between the tobit predictions to the fractional models.

— Inserts Figure 1 about here —

In Models (1)-(6) of Table 3, the variables *age/10*, *country* and *gender* are significant at the 10 percent level. The coefficient for *age/10* measures the effect of increasing age on relative risk aversion, and is found to be positive and significant when housing is included and insignificant when housing is excluded. Household heads born in an English speaking country take fewer financial risks than those born in an non-English speaking country. While *gender* is significant, it is negative when housing is included and positive when housing is dropped as a risky asset. Without housing, *married* and *city* are significant at the 10 percent level but the variable *married* is insignificant with taxes.

Table 4 presents the mean of the average partial effects across all observations in the sample with exogenous wealth. The mean average partial effects appear fairly robust across the model choice. The variable *age/10* is significant for Models (1)-(6) when housing is included, and indicating that being 10 years older will increase the share of risky asset by 0.042-0.045 on average. When housing is excluded as in Models (7)-(12), *age/10* is insignificant. Households with greater than high school education hold a larger share of risky asset by 0.042-0.054 in comparison with households of less education when housing is excluded as a risky asset.

— Insert Table 4 about here —

To address the issue of potential endogenous wealth, Models (1)-(12) in Table 4 were re-estimated using the two stage instrumental variable approach as outlined in section III. On the grounds that wealth is likely to be correlated with that of one's family, a possible instrumental variable for *Wealth* is mother's occupation. Mother's occupation is captured by the dummy variables *mocc1, mocc2, mocc3, mocc4, mocc5* and *mocc6* defined in Table 1. In order for mother's occupation to be a valid instrument, it must be uncorrelated with the error term in the wage equation and correlated with *Wealth*. We assume that *mocc1-6* and error term θ are uncorrelated. To check the latter requirement we estimate the reduced form of *Wealth* by regressing *Wealth* on *mocc1* and all exogenous variables appearing in the equation. Table 5 presents the results to the first stage estimation. The coefficients for *mocc1-6* are positive and significant. Thus, if mother's occupation is uncorrelated with unobserved factors in the error term, we can use it as an instrumental variable for *Wealth*.

— Insert Table 5 about here —

— Insert Table 6 about here —

— Insert Table 7 about here —

Table 6 contains the estimated coefficients used for estimating the APE with endogenous wealth for the fractional probit (FPM) and the double-censored tobit model. The coefficient for the residual obtained from the second stage estimation for the wealth equation, w, is significant for all definitions with the exception of Models (17) and (19). The likelihood ratio test for the inclusion of w confirms endogeneity of wealth for all Tobit models and the FPM Model (13). To determine the direction and size of the bias, we compare the mean average partial effects of wealth in Tables 4 and 7.

Table 7 presents the mean APE with endogenous wealth. The effect of *country*, *married*, *gender*, *city*, *employed*, *retire* and *Wealth* are highest and significant for Models (13a)-(16a) when housing is included, and lower but insignificant when housing is excluded for Models (17a)-(20a). With housing, females on average hold 0.067 to 0.192 less in the share of risky assets to wealth than males. Married individuals on average hold 0.043-0.058 less in the share of risky assets to wealth. Tax-adjustments tend to reduce the effect by around 0.01 except for the variable *employed*, where the effect is larger. Finally, the variable of interest, *Wealth*, is significant at the one percent level for all definitions. The mean APE of wealth with exogenous wealth is larger than for exogenous wealth. This is evidence for a downward bias in Wealth.

— Insert Figure 2 about here —-

Figure 2 presents the graph for the APE averaged over the entire sample for all four alternative definitions for the risky portfolio assuming exogenous wealth (diagram a and b) and endogenous wealth (c and d). Several patterns are observed from Figure 2. First, the mean APE is similar across model specifications. Second, without taxes, the effect of an increase in wealth on the share of risky asset is greater for households with lower wealth at around the \$250,000 to \$500,000. With taxes, the effect of wealth is greatest at around 750,000 to 1 million dollars of wealth for households. Overall, the effect of a change in wealth is less for households with high levels of wealth. Further, after accounting for the endogeneity of wealth, the mean APE of wealth is consistently greater than the mean APE with exogenous wealth, suggesting that the APE with exogeneous wealth is potentially underestimated.

—Insert Table 8 about here —

Table 8 presents the results for the Vuong test for non-nested models. The Vuong test identifys the model that is closer to the true specification. The results from the Vuong test are significant, and therefore rejecting the null hypothesis that there is no difference in between the models. Except for the first definition that includes housing, the fractional models are closer to the true specification than the double-censored tobit model across all definitions under the assumption of strict exogeneity of wealth. Furthermore, under the assumption of endogenous wealth, the fractional probit model is closer to the true specification than the tobit across all definitions.

VI Discussion

This study has analysed the investment of risky assets in a cross-sectional sample of Australian households. Following previous studies, we allow for both a broad and narrow definition of risky assets that includes and that excludes housing wealth as risky, respectively. We also consider the effect of taxes on portfolio holdings as in Bellante and Green (2004).

Our results indicate that the relationship between wealth and the relative risk aversion is quadratic, decreasing relative risk aversion (DRRA) is observed for most households[12] and increasing relative risk aversion (IRRA) observed among the wealthiest households (2.2-2.9

[12]A majority of households in the sample have wealth below 1 million dollars and very few above 1 million dollars.

million dollars and above). A possible explanation for DRRA is that wealthier households under approximately 1.5 million dollars may be better prepared than less wealthy households to sustain the short-term losses with riskier investments that generate higher returns in the long-term. Furthermore, they are more likely to have funds available to afford investment in assets with differing risk and return structure and as a result hold a greater portion of their wealth in risky assets. A possible explanation for the IRRA observed among the wealthiest is that these households prefer to invest most of their wealth in non-risky assets.

This study also tests the hypothesis that females are more risk averse than males. The results indicate that the effect of gender on risky asset holdings is less than definite and differs across definitions of risky assets. Females have greater risk aversion than males when housing is excluded as risky, but males are more risk averse to females when housing is defined as a risky asset. One explanation for this observed result is that females are more likely to invest in housing than other forms of risky assets to males. Moreover, under the narrow definition that excludes housing wealth as risky, gender is not important in explaining portfolio holdings after endogeneity of wealth is accounted for.

This research also accounts for the effect of taxes on the share of risky assets by adopting the framework adopted by Bellante and Green (2004). Results are consistent with the findings by King and Leape (1998) that taxes have little impact in determining the share of risky assets, but they find that taxes are more important in explaining the demand for risky assets. Although this paper improves on measures of aggregate effective tax rate as adopted in papers by Bellante and Saba (and 1986) and Morin and Suarez (1983) by using the average household tax rate,[13] the marginal tax rate for each household is arguably a more adequate measure for capturing the effect of taxes.[14]

Finally, this study accounts for the potential endogeneity of wealth, and find evidence emphasising it's importance in explaining household portfolio holdings. The endogeneity bias is towards zero thereby suggesting past studies may have underreported the effect of wealth. While the findings in this study do not result in substantial differences in the statistical significance of the control variables for the share of risky assets, they indicate that the extent of the magnitude and significance may be misreported if the endogeneity of wealth is not accounted for. At the very least, the influence of wealth and household heterogeneity on the household share of risky assets held in wealth cannot be properly understood without full consideration of endogenous wealth.

Like all such studies, this study too has inevitable limitations. A possible avenue for future

[13]The average household tax rate is measured as the total household taxes over household taxable income.

[14]King and Leape (1998) for example use the marginal tax rate in their analysis. It is the rate of tax on the highest dollar of income you earn in a tax year, and determined at the individual level.

research examining the effects on the share of risky assets to wealth is to investigate a longitudinal data when more waves from the HILDA data on wealth becomes available to capture life-cycle effects and other time varying effects. Unlike most datasets used in the literature, the HILDA survey allows for the identification of the household head but the information is incomplete and the household head is not easily deduced. Finally, as previously discussed the measure for taxes can be improved with calculating the marginal tax rate for each household instead of the average household tax rate.

VII Conclusion

The findings of this paper are quite promising for future research. We find that the FPM and FLM are reasonable models for future studies on portfolio holdings. The essential difference of this research is that we control for the endogeneity of wealth. We obtain that the endogeneity of wealth is important and biases results towards zero. Furthermore, under endogenous wealth, the FPM is recommended over the widely adopted double censored tobit model. A possible avenue for future research is extending longitudinal data using a dynamic FPM as explored by Wooldridge (2005) to allow for life-cycle effects.

References

Arrondel, L. (2000). Risk management and wealth accumulation behavior in france. *economic Letters 74*, 187–194.

Arrow, K. J. (1974). *Essays in the theory of risk-bearing.* Amsterdam: North-Holland Publishing Company.

Bellante, D. and C. Green (2004). Relative risk aversion among the elderly. *Review of Financial Economics 13*(3), 269–281.

Bellante, D. and R. P. Saba (1986). Human capital and life-cycle effects on risk aversion. *The journal of financial research 9*(1), 41–51.

Bodie, Z., R. C. Merton, and W. F. Samuelson (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control 16*(3-4), 427–449.

Davies, J., S. Sandstrom, A. Shorrocks, and E. Wolff (2008). The world distribution of household wealth. *WIDER discussion paper WDP* (3).

Faig, M. and P. Shum (2005). What explains household stock holdings. *Journal of Banking and Finance 30*(9), 2579–2597.

Friend, I. (1973). Mythodology in finance. *The Journal of Finance 28*(2), 257–272.

Gourieroux, C., A. Monfort, and A. Trognon (1984). Pseudo maximum likelihood methods: Theory. *Econometrica 52*(3), 681–700.

Guiso, L., T. Jappeli, and D. Terlizzese (1996). Income risk, borrowing constraints, and portfolio choice. *The American Economic Review 86*(1), 158–172.

Guiso, L. and M. Paiella (2001). Risk aversion, wealth and background risk. SSRN eLibrary.

Hochguertel, S., R. Alessie, and A. van Soest (1997). Saving accounts versus stocks and bonds in household portfolio allocation. *Scandinavian Journal of Economics 99*(1), 81–97.

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

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

King, M. A. and J. I. Leape (1998). Wealth and portfolio composition: Theory and evidence. *Journal of Public Economics 69*, 155–193.

McCarthy, D. (2004). Household portfolio allocation: A review literature. *Presented at the International Collaborative Forum, ESRI, Tokyo, February 2004.*.

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

Palsson, A.-M. (1996). Does the degree of relative risk aversion vary with household characteristics. *Journal of Economic Psychology 17*, 771–787.

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

Riley, W. B. and V. K. Chow (1992). Asset allocation and individual risk aversion. *Financial Analysis Journal 48*(6), 32–7.

Siegel, F. and J. Hoban (1982). Relative risk aversion revisited. *Review of Economics and Statistics 64*(3), 481–487.

Wooldridge, J. (2005). Unobserved heterogeneity and estimation of average partial effects. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Chapter 3. New York: Cambridge University Press.

Appendix

Appendix 1: construction of household head

Respondents are asked to subjectively identify the household decision maker in the Self-Completion Questionnaire of the HILDA survey. The following questions are asked: “Who makes decisions about the following issues in this household?” A. Managing day-to-day spending and paying bills; B. Making large household purchases (e.g. cars and major appliances); C. Savings, investments and borrowing

The respondent must choose from the following set of answers for each question: 1. Me/Mainly me 2. Mainly my spouse/partner; 3. Shared equally between partner and myself ; 4. Someone else (living here or elsewhere); 5. Shared equally among all household members.

For this study, decision-making in savings, investment and borrowing (question C) is considered to be the more appropriate measure for identifying the household head in economic decision-making. The set of responses provided to identify the household head are obvious from those that indicate (1) “me/mainly me” or (2) “Mainly my spouse/partner” as the decision maker. It is possible to directly identify the household head from responses that indicate “me/mainly me”. From responses that indicate the “spouse/partner”, it is possible to indirectly identify the household head by using the information on relationships within the household. However, this study cannot confidently identify the household head from the other responses. Responses (3) and (5) suggest joint decision-making, which may be of interest to study, but beyond the scope of this analysis. Therefore, this analysis only includes households with responses that indicate (1) or (2).

Inconsistencies in reporting household head

The household head variables are subjective measures. This means that there can be inconsistent reporting on who makes decisions in the household. Three types of inconsistencies were encountered and outlined below.

Type 1: Within one household there may be more than one person reporting to be the main decision maker. To account for this type of inconsistency, the following indicator variable was created: *Datahead1* (1,2) where, 1: respondent indicates that the household decision maker as "mainly me" and responses are not inconsistent in the Type 1 sense. 2: inconsistent in the Type 1 sense.

Conversely, there were cases where no one in the same household reported themselves as the household head, but all members indicated another person (spouse/partner) as the decision maker. These observations are also considered inconsistent. For example, Type 2: Mr. X identifies his spouse as the decision maker, but his spouse identifies him as the decision maker. To account for this type of inconsistency, the following indicator variable was created: *Datahead2* (1,2) where, 1: respondent is indicated by their spouse as being the decision maker and responses are not inconsistent in the Type 2 sense. 2: inconsistent in the Type 2 sense.

Finally, this last case refers to inconsistencies observed in direct versus indirect reporting in a household. One would surmise that those that are indicated by their spouses to be the household head, have also indicated themselves as the household decision maker but this is at times not the case. Type 3: Two different examples are posed here. 1) Mr X is indicated by his spouse, but another member in the same household also indicated themselves as the household head; or 2) Mr X is indicated by his spouse, but another member in the household was also indicated as the household head by their spouse. For variables *datahead1* and *datahead2*, this type of inconsistency is essentially the case where datahead1 and/or datahead2 are indicated to be equal to one for more than one household member. To account for this type of inconsistency, the following indicator variable was created: *Datahead3* (1,2) where, 1: respondent indicates that *datahead1*=1 or *datahead2*=1 but responses are not inconsistent in the Type 3 sense. 2: inconsistent in the Type 3 sense.

— Insert Table 9 about here —

Type 3 inconsistencies occur in 34 of the respondents (17 households); further investigation revealed that 94% of these discrepancies were due to parent and child reporting. For example, a particular household has a male 26-year-old respondent indicating himself as the household head, and his mother simultaneously indicates his 59-year-old father as the household decision maker. The father however, indicates makes joint decisions with his partner.

Furthermore, cases where indicator variables *datahead1* and *datahead2* implied inconsistent reporting were also independently identified in the dataset and dropped. However, of the 5 inconsistent responses (Type 2-1), two respondents were identified as the household head to the extent that they belonged to the first household income unit.

Finally, of the 3689 households, those were dropped where the household head did not participate in the Self-completion questionnaire. This reduced the number of observations to 3660.

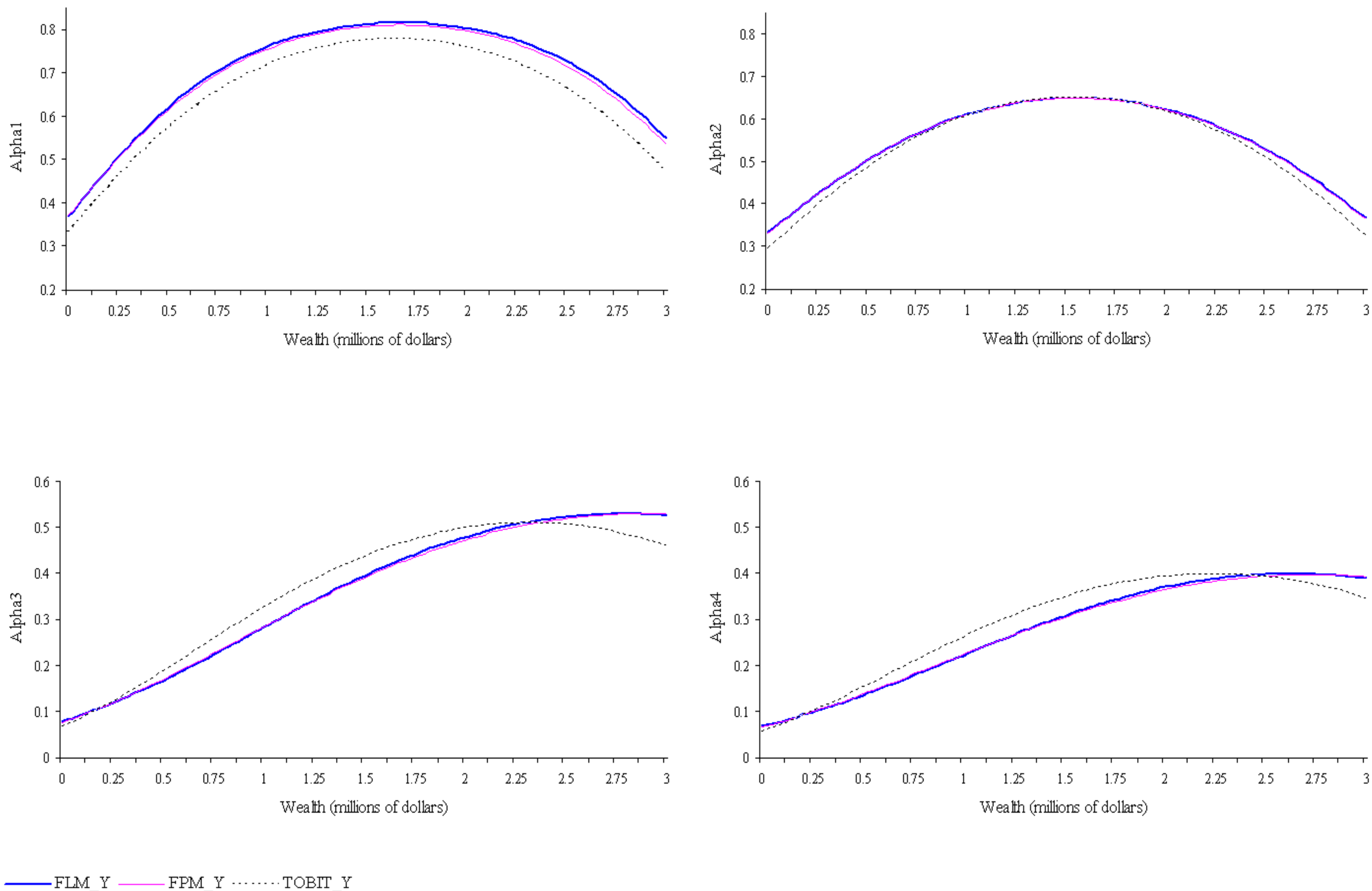
Tables and Figures

Table 1: Variable description (n=2306^)

Variable(s)	Type	Description	Mean	Std. dev	Min	Max	Std. error
age	Continuous	The age of the household head in years.	45.71	15.81	15	90	0.329
wealth	Continuous	net worth in millions of Australian dollars >0	0.39	0.53	1.0E-06	3	0.011
Education	Dummy	highest education greater than high school.	0.73	0.44	0	1	0.0092
Country	Dummy	born in English speaking country	0.91	0.28	0	1	0.0059
Married	Dummy	legally married.	0.34	0.47	0	1	0.0098
gender	Dummy	Household head is female.	0.505	0.5	0	1	0.0104
city	Dummy	Living in the city (ABS standards)	0.61	0.49	0	1	0.0102
employed	Dummy	currently employed and in labour force.	0.69	0.46	0	1	0.0097
retire	Dummy	currently retired from labour force.	0.19	0.39	0	1	0.0082
Share of risky assets**							
Alpha1	Continuous	Risky assets defined to include equity in homes,	0.52	0.35	0	1	-
Alpha2	Continuous	Tax-adjusted* and includes equity in homes	0.44	0.31	0	1	-
Alpha3	Continuous	Risky assets defined to exclude equity in homes	0.16	0.25	0	1	-
Alpha4	Continuous	Tax-adjusted* and excludes equity in homes.	0.13	0.21	0	1	-
Mother's occupation***							
mocc1	Dummy	managers and admin,	0.04	0.2	0	1	0.0043
mocc2	Dummy	professionals and associate professionals,	0.27	0.44	0	1	0.0093
mocc3	Dummy	tradepersons and related workers,	0.08	0.27	0	1	0.0056
mocc4	Dummy	advanced clerical and service workers,	0.09	0.29	0	1	0.00596
mocc5	Dummy	intermediate clerical, sales and service workers;	0.2	0.4	0	1	0.0083
mocc6	Dummy	elementary clerical, sales and service workers;	0.15	0.36	0	1	0.0074

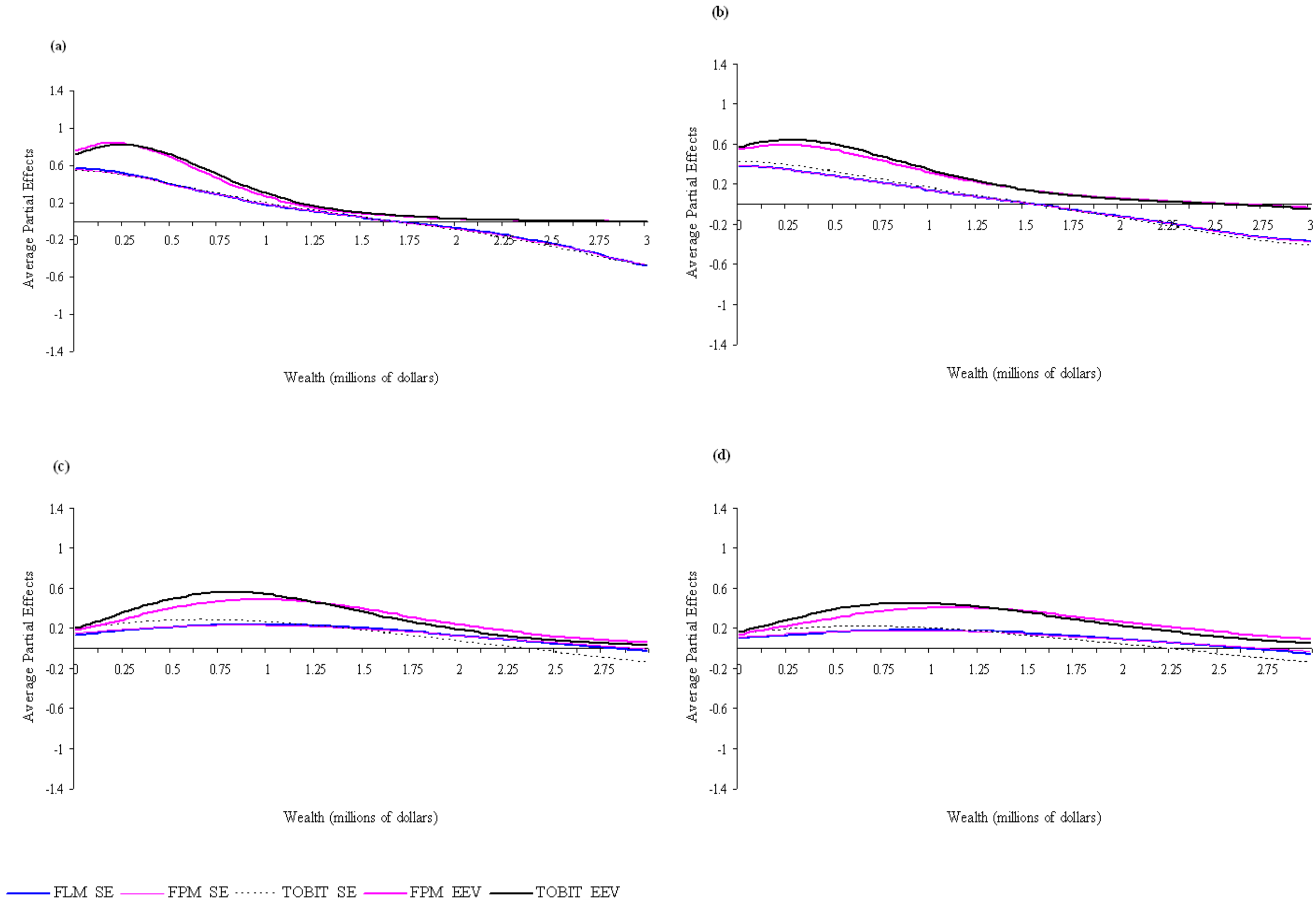
Source: HILDA, Wave2. ^Following the literature, we only consider households with positive wealth. Furthermore, after removing observations where mother's occupation is missing, the number of observations is reduced to 2306.*Tax is calculated as the average household tax rate calculated as the household financial year taxes over household financial year income. **All risky assets include financial equity, business equity and equity in property. ***Reference group are labourers and related workers.

Figure 1: Estimated share of risky assets in wealth vs household wealth



*Source:*HILDA survey, wave2. Wealth is assumed to be strictly exogeneous.

Figure 2: Mean average partial effects of Wealth



Graph of the estimated Average Partial Effect (APE) of wealth averaged across all households on (a) Alpha1: risky assets defined to include equity in home (b) Alpha2: includes equity in home and adjusted for tax (c) Alpha3: excludes equity in home (d) Alpha4: excludes equity in home and adjusted for tax. SE and EEV refers whether wealth is assumed to be "strictly exogenous" or "endogenous explanatory variable", respectively. *Source:* HILDA survey, wave2.

Table 2: Share of risky assets distributed over wealth quantiles

	Wealth quantile*				
	1	2	3	4	Total
alpha1	0.148	0.601	0.678	0.680	0.526
	(0.012)	(0.013)	(0.010)	(0.010)	(0.007)
alpha2	0.131	0.524	0.577	0.536	0.442
	(0.011)	(0.012)	(0.010)	(0.008)	(0.006)
alpha3	0.075	0.119	0.142	0.294	0.157
	(0.009)	(0.010)	(0.010)	(0.011)	(0.005)
alpha4	0.064	0.099	0.117	0.230	0.128
	(0.008)	(0.008)	(0.008)	(0.009)	(0.004)
Financial risk aversion	0.960	0.910	0.913	0.790	0.893
	(0.004)	(0.005)	(0.07)	(0.08)	(0.006)

Source: HILDA survey, wave 2. Note: Standard errors of the mean are reported in the brackets. Financial risk aversion is dummy variable = 1 if the individual is unwilling to take substantial risks in order to have substantial gains, 0 otherwise. *First quantile value for wealth is $58,035; second quantile at $212,000; third quntile at $509,600; and fourth quantile for wealth at $3,000,000. Number of observations (n) =2306.

Table 3: Estimation results: strictly exogenous wealth

Dependent variable (Y)	*Alpha1*	*Alpha1*	*Alpha1*	*Alpha2*	*Alpha2*	*Alpha2*	*Alpha3*	*Alpha3*	*Alpha3*	*Alpha4*	*Alpha4*	*Alpha4*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables	FLM	FPM	TOBIT	FLM	FPM	TOBIT	FLM	FPM	TOBIT	FLM	FPM	TOBIT
constant	-1.334***	-0.829***	-0.014	-1.339***	-0.842***	0.009	-2.654***	-1.545***	-0.321***	-2.742***	-1.581***	-0.25***
	(0.1887)	(0.1647)	(0.039)	(0.1708)	(0.1043)	(0.0404)	(0.2548)	(0.1391)	(0.0511)	(0.2503)	(0.1333)	(0.0463)
age/10	0.189***	0.117***	0.058***	0.194***	0.12***	0.058***	-0.028	-0.016	-0.002	-0.013	-0.009	0
	(0.0275)	(0.0251)	(0.0071)	(0.0244)	(0.015)	(0.0064)	(0.0376)	(0.0207)	(0.008)	(0.0364)	(0.0197)	(0.0065)
Education	0.071	0.046	0.035*	0.043	0.027	0.025	0.461***	0.254***	0.108***	0.433***	0.236***	0.088***
	(0.0723)	(0.0659)	(0.0203)	(0.0656)	(0.0404)	(0.0181)	(0.1032)	(0.0548)	(0.0213)	(0.1025)	(0.0532)	(0.0174)
Country	-0.194**	-0.121*	-0.048*	-0.187**	-0.114**	-0.043*	-0.116	-0.057	-0.006	-0.103	-0.049	-0.004
	(0.1026)	(0.0962)	(0.0285)	(0.0886)	(0.0546)	(0.0252)	(0.1489)	(0.0835)	(0.0227)	(0.1436)	(0.0786)	(0.0262)
Married	0.129**	0.079**	0.05**	0.089*	0.055*	0.037**	-0.242***	-0.129***	-0.027	-0.217***	-0.115**	-0.019
	(0.0626)	(0.0631)	(0.0196)	(0.0531)	(0.0329)	(0.0168)	(0.0887)	(0.0497)	(0.019)	(0.0833)	(0.0456)	(0.0156)
Gender	-0.204***	-0.127***	-0.056***	-0.227***	-0.139***	-0.062***	0.222***	0.118**	0.037**	0.193	0.1**	0.027*
	(0.0607)	(0.0567)	(0.0179)	(0.0532)	(0.0328)	(0.0154)	(0.0837)	(0.0461)	(0.0176)	(0.0809)	(0.0435)	(0.0145)
City	-0.02	-0.011	-0.012	-0.09**	-0.055	-0.029*	-0.219***	-0.119***	-0.032*	-0.26***	-0.138***	-0.032
	(0.0618)	(0.0568)	(0.0173)	(0.0548)	(0.0337)	(0.0152)	(0.0832)	(0.0461)	(0.0176)	(0.0801)	(0.0436)	(0.0146)
employed	0.196*	0.127**	0.101***	0.022	0.019	0.044*	0.335**	0.186**	0.12***	0.194	0.109	0.081***
	(0.1009)	(0.0864)	(0.0278)	(0.0965)	(0.0591)	(0.0241)	(0.16)	(0.0844)	(0.0294)	(0.1616)	(0.0838)	(0.024)
retire	0.015	0.016	0.026	0.126	0.083	0.047	-0.01	-0.002	0.031	0.013	0.013	0.027
	(0.1346)	(0.1199)	(0.0392)	(0.1246)	(0.0767)	(0.033)	(0.1999)	(0.1071)	(0.0398)	(0.1983)	(0.1043)	(0.0317)
Wealth2	-0.751***	-0.453***	-0.226***	-0.569***	-0.354***	-0.186***	-0.332***	-0.18***	-0.134***	-0.32***	-0.17***	-0.11***
	(0.0587)	(0.055)	(0.0174)	(0.0482)	(0.0299)	(0.0148)	(0.0597)	(0.0351)	(0.0163)	(0.0541)	(0.0313)	(0.0133)
Wealth	2.508***	1.502***	0.745***	1.76***	1.091***	0.569***	1.85***	1.044***	0.616***	1.672***	0.918***	0.487***
	(0.1572)	(0.1398)	(0.0445)	(0.1264)	(0.0779)	(0.038)	(0.173)	(0.0994)	(0.0432)	(0.1611)	(0.0908)	(0.0351)
sigma			0.396***			0.342***			0.359***			0.294***
			(0.0072)			(0.006)			(0.0078)			(0.0063)
Assumption on *Wealth*	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE
*Wealth*at $\frac{dY}{dWealth}=0$	1.68	1.66	1.648	1.546	1.541	1.53	2.753	2.9	2.299	2.612	2.7	2.213
Log-likelihood	-1457.9	-1459.5	-1415.1	-1489.6	-1489.2	-1105.6	-920.27	-919.35	-1134.8	-826.04	-825.3	-854.86
AIC	2937.8	2941	2852.2	3001.2	3000.4	2233.2	1862.54	1860.7	2291.6	1674.08	1672.6	1731.72

*Source:*HILDA survey, wave 2. Note: ***,** , *and indicate significance at the one, 5 and 10% levels respectively. FLM, FPM and TOBIT denote the fractional logit estimator, the fractional probit estimator, and the double-censored tobit (censored at 1 and 0), respectively. SE and EEV stand for "strictly exogenous" and "endogenous explanatory variables", respectively. The quantities in (.) below estimates are the standard errors. For the fractional logit and fractional probit estimators, the standard errors computed are the Papke and Wooldridge's (1996) standard errors that are robust to variance misspecification. AIC, BIC and CAIC refer to the Akaike, Bayesian and the Constant Akaike information criterion, respectively. Number of observations (n) =2306.

Table 4: Estimation Results: Mean average partial effects with strictly exogenous wealth

Dependent variable	*Alpha1*	*Alpha1*	*Alpha1*	*Alpha2*	*Alpha2*	*Alpha2*	*Alpha3*	*Alpha3*	*Alpha3*	*Alpha4*	*Alpha4*	*Alpha4*
	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)	(7a)	(8a)	(9a)	(10a)	(11a)	(12a)
	FLM	FPM	TOBIT	FLM	FPM	TOBIT	FLM	FPM	TOBIT	FLM	FPM	TOBIT
age/10	0.042***	0.043***	0.042***	0.044***	0.044***	0.045***	-0.003	-0.004	-0.001	-0.001	-0.002	9.50E-05
	(0.006)	(0.006)	(0.0052)	(0.0054)	(0.0055)	(0.0051)	(0.0046)	(0.0045)	(0.0039)	(0.0038)	(0.0038)	(0.0032)
education	0.016	0.017	0.025*	0.01	0.01	0.02	0.052***	0.053***	0.05***	0.042***	0.043***	0.041***
	(0.016)	(0.0161)	(0.0147)	(0.0149)	(0.0149)	(0.0141)	(0.0109)	(0.0108)	(0.0094)	(0.0093)	(0.0092)	(0.0078)
country	-0.043*	-0.044*	-0.035*	-0.043**	-0.042**	-0.034*	-0.015	-0.013	-0.003	-0.011	-0.01	-0.002
	(0.0225)	(0.0228)	(0.0206)	(0.0204)	(0.0206)	(0.02)	(0.0191)	(0.0192)	(0.0112)	(0.016)	(0.016)	(0.0131)
married	0.029**	0.029**	0.036**	0.02**	0.02*	0.029**	-0.029***	-0.028***	-0.013	-0.022***	-0.022***	-0.009
	(0.014)	(0.0139)	(0.0144)	(0.0122)	(0.0121)	(0.0133)	(0.0103)	(0.0104)	(0.0091)	(0.0084)	(0.0086)	(0.0076)
gender	-0.045***	-0.046***	-0.04***	-0.052***	-0.051***	-0.048***	0.027***	0.026**	0.018**	0.02**	0.019**	0.014*
	(0.0133)	(0.0134)	(0.0129)	(0.012)	(0.0121)	(0.012)	(0.0101)	(0.0101)	(0.0086)	(0.0084)	(0.0085)	(0.0072)
city	-0.004	-0.004	-0.008	-0.021*	-0.02	-0.022*	-0.027***	-0.026**	-0.016*	-0.028***	-0.027***	-0.016**
	(0.0136)	(0.0137)	(0.0125)	(0.0124)	(0.0125)	(0.0119)	(0.0104)	(0.0104)	(0.0087)	(0.0088)	(0.0088)	(0.0074)
employed	0.043*	0.046**	0.073***	0.005	0.007	0.034*	0.039**	0.04**	0.056***	0.02	0.021	0.039***
	(0.0221)	(0.0226)	(0.0198)	(0.0219)	(0.0221)	(0.0186)	(0.0178)	(0.0172)	(0.0131)	(0.016)	(0.0156)	(0.011)
retire	0.003	0.006	0.019	0.029	0.031	0.037	-0.001	0.0003	0.015	0.001	0.003	0.014
	(0.0298)	(0.03)	(0.0285)	(0.0288)	(0.0289)	(0.0264)	(0.0242)	(0.0233)	(0.0201)	(0.021)	(0.0203)	(0.0164)
wealth	0.438***	0.426***	0.417***	0.296***	0.299***	0.329***	0.178***	0.186***	0.23***	0.136***	0.143***	0.182***
	(0.0287)	(0.0279)	(0.0294)	(0.0237)	(0.0238)	(0.0266)	(0.0158)	(0.0168)	(0.0197)	(0.0127)	(0.0136)	(0.0161)
Assumption on *Wealth*	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE	SE

*Source:*HILDA survey, wave 2. Note: ***,**,* and indicate significance at the 1, 5 and 10% levels respectively. FLM, FPM and TOBIT denote the fractional logit estimator, the fractional probit estimator, and the double-censored tobit (censored at 1 and 0), respectively. SE and EEV stand for "strictly exogenous" and "endogenous explanatory variables", respectively. The quantities in (.) below estimates are the standard errors, estimated using the delta method. Number of observations (n) =2306.

Table 5: Estimation Results: first stage estimation (Dependent variable = *Wealth)*

variable	coefficient
constant	0.586***
	(0.625)
age/10	0.114***
	(0.009)
Education	0.145***
	(0.024)
Country	0.039
	(0.035)
Married	0.268***
	(0.022)
Gender	0.089***
	(0.021)
City	0.071***
	(0.021)
employed	0.101***
	(0.031)
retire	-0.067
	(0.044)
mocc1	0.203***
	(0.054)
mocc2	0.085***
	(0.031)
mocc3	0.113***
	(0.113)
mocc4	0.133***
	(0.133)
mocc5	0.078**
	(0.078)
mocc6	0.749**
	(0.075)
observations	2306
R-squared	0.201

Source: HILDA survey, wave 2 Note: ***, **, and * indicate significance at the 1, 5 and 10% levels respectively. The quantities in (.) below estimates are the least squares standard errors.

F-statistic for test of $\delta_{mocc1} = \delta_{mocc2} = \delta_{mocc3} = \delta_{mocc4} = \delta_{mocc5} = \delta_{mocc6} = 0$ is 3.55 with p-value=0.0017.

Table 6: Estimation Results: endogeneity test

Dependent variable (Y)	Alpha1		Alpha2		Alpha3		Alpha4	
	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Variables	FPM	TOBIT	FPM	TOBIT	FPM	TOBIT	FPM	TOBIT
Constant	-0.307	0.26	-0.481***	0.188***	-1.167***	-0.131	-1.258***	-0.108
	(0.2271)	(0.2156)	(0.2021)	(0.0943)	(0.2805)	(0.1087)	(0.2627)	(0.0852)
age/10	0.002	-0.003	0.04	0.018	-0.1*	-0.044***	-0.08	-0.031**
	(0.0478)	(0.0459)	(0.0427)	(0.0194)	(0.0624)	(0.0222)	(0.0584)	(0.0178)
education	-0.111*	-0.047	-0.081	-0.029	0.141*	0.052*	0.14*	0.046*
	(0.0738)	(0.0643)	(0.0663)	(0.031)	(0.0913)	(0.0354)	(0.0874)	(0.028)
country of birth	-0.162***	-0.069**	-0.141***	-0.056***	-0.088	-0.02	-0.075	-0.015
	(0.0642)	(0.0365)	(0.056)	(0.027)	(0.0832)	(0.0317)	(0.0781)	(0.0273)
married	-0.191**	-0.092	-0.132	-0.056	-0.326***	-0.125***	-0.283***	-0.092***
	(0.1105)	(0.1009)	(0.0983)	(0.0457)	(0.1457)	(0.0528)	(0.1366)	(0.0426)
gender	-0.218***	-0.104***	-0.203***	-0.093***	0.051	0.003	0.043	0.003
	(0.0519)	(0.0375)	(0.0462)	(0.0211)	(0.0675)	(0.0219)	(0.0637)	(0.0179)
city	-0.084**	-0.05*	-0.105***	-0.054***	-0.172***	-0.059***	-0.182***	-0.052***
	(0.0458)	(0.0328)	(0.0409)	(0.0194)	(0.0578)	(0.022)	(0.0547)	(0.0179)
employed	0.024	0.047	-0.053	0.008	0.111	0.082***	0.045	0.052**
	(0.0723)	(0.0458)	(0.0677)	(0.0285)	(0.0976)	(0.0347)	(0.0951)	(0.0278)
retire	0.088	0.064	0.133*	0.072***	0.051	0.058	0.057	0.047
	(0.0877)	(0.0524)	(0.0819)	(0.0349)	(0.1138)	(0.0414)	(0.1104)	(0.0333)
wealth2	-0.451***	-0.226***	-0.353***	-0.185***	-0.18***	-0.134***	-0.169***	-0.11***
	(0.0356)	(0.0174)	(0.0298)	(0.0148)	(0.0351)	(0.0162)	(0.0313)	(0.0132)
wealth	2.51***	1.273***	1.79***	0.915***	1.778***	0.984***	1.544***	0.761***
	(0.4023)	(0.3755)	(0.3582)	(0.1636)	(0.5268)	(0.1878)	(0.4957)	(0.1508)
w	-1.019***	-0.534*	-0.708***	-0.35***	-0.742	-0.372***	-0.634	-0.277**
	(0.3962)	(0.3703)	(0.3538)	(0.1605)	(0.5231)	(0.1842)	(0.4918)	(0.1481)
sigma		0.396***		0.342***		0.359***		0.293***
		(0.0072)		(0.006)		(0.0078)		(0.0063)
Assumption on *Wealth*	EEV	EEV	EEV	EEV	EEV	EEV	EEV	EEV
*Wealth*at $\frac{dY}{dWealth} = 0$	2.78	2.82	2.54	2.47	4.94	3.67	4.56	3.46
Log-likelihood	-1458	-1411.05	-1488.5	-1103.21	-918.84	-1132.77	-824.96	-853.208
AIC	2892	2848.104	3001	2232.427	1861.68	2291.546	1673.92	1732.416
LRT	-3**	-8.096***	-1.4	-4.77***	-1.02	-4.05***	-0.68	-3.3**

*Source:*HILDA survey, wave 2. Note: ***,** and * indicate significance at the one, 5 and 10% levels respectively. EEV stands "endogenous explanatory variables". The quantities in (.) below estimates are the standard errors. For the FPM estimators, the standard errors computed are the Papke and Wooldridge's (1996) standard errors that are robust to variance misspecification. LRT is the likelihood ratio test for endogenous wealth. Number of observations (n) =2306.

Table 7: Estimation Results: Mean average partial effects with endogenous wealth

Dependent variable	*Alpha1*	*Alpha1*	*Alpha2*	*Alpha2*	*Alpha3*	*Alpha3*	*Alpha4*	*Alpha4*
	(13a)	(14a)	(15a)	(16a)	(17a)	(18a)	(19a)	(20a)
Variables	FPM	TOBIT	FPM	TOBIT	FPM	TOBIT	FPM	TOBIT
age/10	0.001	-0.0005	0.013	0.011	-0.018*	-0.018***	-0.015*	-0.014**
	(0.0156)	(0.0136)	(0.0137)	(0.0139)	(0.0115)	(0.0087)	(0.0103)	(0.0081)
education	-0.034*	-0.026	-0.027	-0.02	0.033**	0.025**	0.025**	0.021*
	(0.0209)	(0.0212)	(0.0191)	(0.0186)	(0.0191)	(0.0146)	(0.0152)	(0.013)
country	-0.05***	-0.069***	-0.047***	-0.061***	1.70E-02	0.017	-0.014	-0.007
	(0.0207)	(0.032)	(0.0183)	(0.0281)	(0.0261)	(0.0203)	(0.0162)	(0.0121)
married	-0.058***	-0.125***	-0.043**	-0.105***	-0.047	-0.035	-0.05***	-0.041***
	(0.0288)	(0.0503)	(0.0253)	(0.0477)	(0.0435)	(0.0345)	(0.0209)	(0.0169)
gender	-0.067***	-0.192***	-0.067***	-0.173***	-0.042	-0.037	0.008	0.001
	(0.0131)	(0.0532)	(0.013)	(0.0531)	(0.0528)	(0.0416)	(0.012)	(0.0103)
city	-0.026**	-0.218***	-0.035***	-0.208***	-0.073	-0.059	-0.035***	-0.024***
	(0.0151)	(0.0576)	(0.0125)	(0.0586)	(0.0594)	(0.0478)	(0.0102)	(0.0085)
employed	0.007	-0.196***	-0.018	-0.208***	-0.057	-0.03	0.008	0.024**
	(0.0242)	(0.0714)	(0.0233)	(0.0711)	(0.0698)	(0.0557)	(0.0168)	(0.0141)
retire	0.027	-0.159***	0.045**	-0.159***	-0.048	-0.007	0.011	0.022
	(0.0267)	(0.0754)	(0.0265)	(0.0784)	(0.0705)	(0.0582)	(0.0204)	(0.0161)
wealth	0.7***	0.69***	0.515***	0.557***	0.297***	0.35***	0.253***	0.296***
	(0.0841)	(0.0761)	(0.0805)	(0.0835)	(0.0881)	(0.0659)	(0.0819)	(0.0628)
Assumption on *Wealth*	EEV	EEV	EEV	EEV	EEV	EEV	EEV	EEV

Source: HILDA survey, wave 2 Note: ***, **, and * indicate significance at the 1, 5 and 10% levels respectively. FPM and TOBIT denote the fractional probit estimator, and the double-censored tobit (censored at 1 and 0), respectively. EEV stands for "endogenous explanatory variables". The quantities in (.) below estimates are the standard errors, bootstrapped at 200 bootstraps. Number of observations (n) =2306.

Table 8: Estimation results: Vuong test

Dependent variable	F	G	Assumption on wealth	Vuong statistic*	Preferred model
	FPM(2)	Tobit(3)	SE	-3.58 (0.83)	Tobit
Alpha1	FPM(2)	FLM(1)	SE	-4.93 (0.28)	FLM
	FLM(1)	Tobit(3)	SE	-2.99 (0.84)	Tobit
	FPM(13)	Tobit(14)	EEV	12.66 (47.12)	FPM
	FPM(5)	Tobit(6)	SE	18.56 (0.53)	FPM
Alpha2	FPM(5)	FLM(4)	SE	3.8 (0.05)	FPM
	FLM(4)	Tobit(6)	SE	18.35 (0.53)	FLM
	FPM(15)	Tobit(16)	EEV	18.56 (0.28)	FPM
	FPM(8)	Tobit(9)	SE	5.36 (0.83)	FPM
Alpha3	FPM(8)	FLM(7)	SE	-2.46 (0.08)	FLM
	FLM(7)	Tobit(9)	SE	5.27 (0.83)	FLM
	FPM(17)	Tobit(18)	EEV	34.56 (0.47)	FPM
	FPM(11)	Tobit(12)	SE	42.55 (0.62)	FPM
Alpha4	FPM(11)	FLM(10)	SE	3.028 (0.07)	FPM
	FLM(10)	Tobit(12)	SE	42.56 (0.62)	FLM
	FPM(19)	Tobit(20)	EEV	42.58 (0.38)	FPM

Source: HILDA survey, wave2 Note: Values in (.) in column 2 refer to the associated estimated model number. SE and EEV stand for "strict exogenous" and "endogenous explanatory variables", respectively. The quantities in (.) below estimates are the standard errors. *All estimates are significant at the 0.02 percent level. Number of observations (n) =2306. The Vuong test is a non-nested test that compares two rival models F and G and identifies the model that is closer to the true specification. Let f and g be the log-likehoods at each observation for models F and G, respectively. The null hypothesis is given by H_0 :$\mathrm{E}^0\left(\ln \frac{f(.)}{g(.)}\right) = 0$, meaning the two rival models are equivalent in terms of being close to the true specification. There are two alternative hypothesis. The first is H_f :$\mathrm{E}^0\left(\ln \frac{f(.)}{g(.)}\right) > 0$ meaning model F is favoured to model G. The second alternative hypothesis H_g :$\mathrm{E}^0\left(\ln \frac{f(.)}{g(.)}\right) < 0$, meaning F is not favoured to G. If $|Vuong| < c$, where c is the critical value from the standard normal, then we conclude that there is no discrimination between the two rival models in terms of being close to true specification. If $Vuong > c$, then the null is rejected in favour of H_f, if $Vuong < -c$ then the null is rejected in favour of H_g.

Table 9: Indicators for household head on respondents

	1	2	Total cases
Datahead1	3276	455	3731
		(224)	(3500)
Datahead2	889	68	957
		(34)	(923)
Datahead3	3689	34	3723
		(17)	(3723)

Note: Values in brackets refer to the number of households these responses relate to.