Ownership of Stocks and Mutual Funds: A Panel Data Analysis*

Rob Alessie
Utrecht University & Tinbergen Institute

Stefan Hochguertel
Free University Amsterdam

Arthur van Soest†
Tilburg University

This version: January 22, 2003

Abstract

This paper analyzes the ownership dynamics of stocks and mutual funds, using representative household panel data, the Dutch CentER Savings Survey 1993–1998. A bivariate dynamic binary choice model is introduced, accounting for interactions between the two types of assets. We find that unobserved heterogeneity and state dependence play a large role for both types of assets. The positive relation between ownership of one type in one period and the other type in the next period is explained by correlated unobserved heterogeneity. A negative state dependence effect of lagged ownership of stocks on ownership of mutual funds is found, which can be explained by the costs of shifting funds across the two forms of stock holding.

Keywords: household portfolio choice, panel data

JEL classification: *C33, C35, D12, D91*

*This paper has greatly benefited from useful comments by an anonymous referee, Richard Blundell, other participants of the TMR Conference on Savings, Pensions, and Portfolio Choice, Paris, the 2001 Annual Meeting of the European Economic Association, Lausanne, the 2002 Royal Economic Society Meeting in Warwick, and seminar participants at the European University Institute, Free University Amsterdam, Tilburg University, University of Groningen and University of Cyprus. Financial support from the TMR network on "Savings and Pensions" (grant number: FMRXCT960016) is gratefully acknowledged.

†Correspondence to: Arthur van Soest, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, e–mail: a.h.o.vansoest@uvt.nl

1 Introduction

In many industrialized countries including the Netherlands, the percentage of private households that own some type of risky financial assets has increased substantially during the nineties. In the US for example, the fraction of households owning some risky financial assets increased from 31.9% in 1989 to 49.2% in 1998. In Italy, the ownership rate increased from 12.0% to 22.1% in the same time period.[1] Similar trends exist in many other countries. To quote The Economist of March, 2001: "Wider share ownership is profoundly important." It spreads wealth, changes attitudes to economic freedom and lowering business taxes, and leads to greater shareholder activism. This puts pressure on managers to improve their performance and promises to raise productivity and economic growth. Household stock ownership becomes more and more important with all kinds of implications for financial markets and macro-economic policy. According to the Financial Times of August 30 2000, the wider share ownership has reversed the public opinion on the US Federal Reserve's policy of cutting interest rates: while in the past, the majority of the public would be concerned about lower returns to their savings accounts, most households will now applaud an interest rate cut since it increases the expected returns to their shares portfolio. On the other hand, the same Financial Times article states, referring to the group of retail investors in risky assets, that "one problem for policy makers analyzing this growing group of Americans is that useful data on the identity of the average investor is hard to come by." This illustrates the need for empirical work on portfolio choice at the level of the individual households.

Guiso *et al.* (2002) provide an overview of the current state of the art in this field. This volume links portfolio choice theory to empirical research and contains empirical studies for several countries. While many countries have some survey data on ownership and amounts invested for several types of assets, this data is often limited to one or more cross–sections. Though useful for many purposes, such data is insufficient to analyze the dynamics of portfolio choice behavior. This requires panel data. Household panels with information on portfolio composition are currently available for the US, Italy, and the

[1]These numbers are taken from Guiso *et al.* (2002), Table 3.

Netherlands. Studies that look at ownership dynamics are rare. Alessie *et al.* (2002) present some univariate results for risky financial assets and employer sponsored saving plans, using six annual waves of Dutch data. Ioannides (1992) uses two waves of the US Survey of Consumer Finances to study ownership and portfolio share dynamics of many types of assets including stocks. Both studies find an important effect of lagged ownership on current ownership of the same asset.

According to most models of portfolio choice, all agents hold risk free financial assets, and this is also what is typically seen in the survey data. We therefore focus on risky financial assets. Existing empirical studies typically consider broadly defined asset groups, with all risky financial assets as one category. Important differences between various risky financial assets, however, will not be revealed in an analysis at this high level of aggregation. Although it is infeasible to use survey data to analyze ownership of every single financial product in the market, it seems worthwhile to distinguish a few subcategories of risky financial assets and to investigate the dynamics in the ownership patterns of these categories as well as the interactions between these patterns.

In particular, we think it is useful to consider the two largest categories, individual stocks and mutual funds. The theoretical argument to treat these separately is the dramatic difference in the degree of riskiness and possibly in terms of costs and information "intensiveness", as in King and Leape (1987). At least in theory, a well-diversified mutual fund should be placed on the efficient frontier and provide a certain expected return at minimum risk. Mutual funds thus seem very attractive for the small, non–expert investor who wants to invest a limited amount with relatively low transaction costs. Investment in individual stocks is much riskier but can be motivated on the basis of private, subjective, distributions of future returns or hedging against individual income risk. Moreover, since transaction costs for stocks will be less than proportional with the amounts held, holding individual stocks may be more attractive for the large investors. Portfolio heterogeneity then arises because of differences in observable (e.g., tax rate, socio–economic characteristics) and unobservable (e.g., risk attitude, beliefs about distribution of asset returns) investor characteristics. An empirical argument to distinguish between the two types of risky assets is that in many countries including the Netherlands, the mutual funds market has grown even more than the market for individual stocks.

In this paper, we use dynamic binary choice panel data models to explain the dynamics of the ownership structure of asset portfolios. The existing dynamic random effects probit model for panel data models is extended to the bivariate case, accounting for interactions between two types of assets. One of the main features of the univariate dynamic binary choice model with random effects is that it can distinguish between unobserved heterogeneity and genuine state dependence. In addition, the bivariate model can explain the correlation between ownership of one type of asset and lagged ownership of the other type of asset both from correlated unobserved heterogeneity and from state dependence across assets. The correlation between random effects in the ownership equations captures correlated unobserved heterogeneity. Dummies for lagged ownership of each asset type in each equation capture genuine state dependence effects. To investigate the sensitivity of the results for the random effects assumption, we also compare our model with a fixed effects dynamic linear probability model.

The empirical analysis considers ownership of stocks and mutual funds, using the 1993–1998 waves of the CentER Savings panel survey of Dutch households. This is one of the few existing household panel surveys with detailed information on ownership of many types of assets and debts. The sample consists of a sub–sample designed to be representative for the Dutch population, and of a (smaller) sub–sample from the highest income decile. Since ownership of risky assets is much more common among the rich than among others, this makes the data particularly useful for our purposes. The estimation sample is an unbalanced panel with 2861 households who, on average, participate in 3.4 waves.

Our aim is to increase insight in how households adjust the structure of their asset portfolios, addressing questions such as the following. Who are the people who have invested in mutual funds or stocks? Do background variables such as income, age, education level, and labor market status affect ownership rates of the two types of assets in the same way? Can changes in these background variables explain the increasing trends in the ownership rates? Why has the ownership rate of individual stocks increased less than the ownership rate of mutual funds? Have most new investors gone into mutual funds, or have people replaced individual stocks by mutual funds? If people hold mutual funds to diversify their risk, there seems no reason to hold individual stocks in addition. Still, the raw data show a positive correlation between ownership of mutual funds and ownership of individual stocks. Is this spurious correlation, or is there genuine state dependence across asset types, which could, for instance, be due to learning effects? Or is it because the new mutual funds owners simply keep their individual stocks?

The remainder of this paper is organized as follows. In the next section, the dynamic bivariate probit model is presented. The data are described in Section 3. Section 4 contains estimation results for the probit model and a brief comparison with the results of a fixed effects linear probability model. Section 5 concludes.

2 A Bivariate Dynamic Random Effects Probit Model

In this subsection we introduce a multivariate discrete choice model for panel data, to explain ownership of different types of assets. For the sake of notational convenience, we present the bivariate case, but the generalization to the case of more than two asset types is straightforward. This model explicitly incorporates the binary nature of the dependent variables and produces predicted ownership probabilities between zero and one. It relies on individual effects being uncorrelated with regressors. Since this assumption is hard to relax in the current framework, we will, following Hyslop (1999), compare the results with those of a fixed effects linear probability model (LPM). The LPM has the potential drawback that predicted ownership probabilities may be outside the zero/one interval.[2]

Since the model will be applied to ownership of stocks and mutual funds, asset type 1 is referred to as stocks and asset type 2 as mutual funds. The following notation is used, where the index for the household is suppressed.

y_{jt}: dependent variables; ownership dummies for stocks ($y_{1t} = 1$ if the household owns stocks in year t, $y_{1t} = 0$ otherwise) and mutual funds ($y_{2t} = 1$ if the household owns mutual funds in year t, $y_{2t} = 0$ otherwise); $t = 1, \ldots, T$.

[2] More details on the models and the estimation procedure are given the working paper version of this paper, Alessie *et al.* (2001).

$\boldsymbol{x}_t$: vector of independent variables, assumed to be strictly exogenous. The same independent variables are used in the two ownership equations.

α_j: random individual effects ($j = 1, 2$); (α_1, α_2) is assumed to be bivariate normal with variances $\sigma^2_{\alpha_1}$ and $\sigma^2_{\alpha_2}$ and covariance $\sigma_{\alpha_1}\sigma_{\alpha_2}\rho_\alpha$.

u_{jt}: error terms ($j = 1, 2;\ t = 1, \ldots, T$); (u_{1t}, u_{2t}) are assumed to be independent over time and bivariate standard normal with covariance ρ.

We assume that (α_1, α_2), $\{u_{jt};\ j = 1, 2;\ t = 1, \ldots, T\}$ and $\{\boldsymbol{x}_t;\ t = 1, \ldots, T\}$ are independent (implying that $\boldsymbol{x}_t$ is strictly exogenous).

The following specification will be used.

$$y^\star_{1t} = \boldsymbol{x}'_t\beta_1 + y_{1,t-1}\gamma_{11} + y_{2,t-1}\gamma_{12} + \alpha_1 + u_{1t} \tag{1}$$

$$y^\star_{2t} = \boldsymbol{x}'_t\beta_2 + y_{1,t-1}\gamma_{21} + y_{2,t-1}\gamma_{22} + \alpha_2 + u_{2t} \tag{2}$$

$$y_{jt} = \begin{cases} 1 & \text{if } y^\star_{jt} > 0 \\ 0 & \text{else} \end{cases} \qquad j = 1, 2;\ t = 1, \ldots, T \tag{3}$$

Some special cases are worth mentioning. If $\gamma_{12} = 0$, the equation for stocks (1) does not contain the lagged mutual funds ownership dummy. In that case, the parameters β_1, γ_{11} and $\sigma^2_{\alpha_1}$ can be estimated consistently by considering only equation (1). This would be the standard univariate panel data probit model with state dependence ($y_{1,t-1}$ is included) as well as unobserved heterogeneity (the random effect α_1). See Heckman (1981a) for a discussion of this model. Similarly, the equation for mutual funds (2) can be estimated as a univariate model if $\gamma_{21} = 0$.

If $y_{2,t-1}$ enters the first equation but error terms and random effects in the first equation are independent of error terms and random effects in the second equation, then $y_{2,t-1}$ is weakly exogenous in the equation for y_{1t}. In this case the first equation could be treated as a univariate model with (weakly) exogenous regressors only.

One of the main advantages of dynamic random effects models is their ability to distinguish between unobserved heterogeneity (random effects) and state dependence (the lagged dependent variable). Both phenomena can explain why ownership of stocks in period t is positively correlated with ownership of stocks in period $t + 1$ (conditional on observed background variables $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$). The bivariate extension can in addition address "spill–over effects" from one asset type on the other. If ownership of stocks in period $t+1$ is correlated to ownership of mutual funds in period t, this can be due to correlated unobserved heterogeneity (i.e., a non–zero covariance between α_1 and α_2) or due to state dependence across asset types, i.e., a non–zero value of γ_{12}. This is important for understanding the dynamics of the asset ownership decisions. For example, a positive value of γ_{12} could mean that mutual funds – which are easily accessible and advertised on a large scale – may have a learning effect in the sense that their acquisition changes people's attitudes to holding risky assets in general. People may then be induced to start buying individual stocks. On the other hand, a positive correlation between the random effects would simply mean that the same people who find it attractive to hold stocks in general also have a preference for holding mutual funds.

Some of the restrictive features of the specification in (3) are easy to generalize. We have estimated univariate specifications allowing for first order autocorrelation in u_{jt} but found insignificant values of the autocorrelation coefficient for both assets. Adding interactions of the two lagged dependent variables or of lagged dependent variables with $\boldsymbol{x}_t$ would make the model as flexible as a transition model with four different ownership states (both assets owned, stocks only, mutual funds only, neither of the two; the standard transition model would not include the random effects, however). We experimented with interaction terms but found they were mostly insignificant and did not change the qualitative conclusions. The assumption that α_1 and α_2 are independent of the time varying regressors can be relaxed by including linear combinations of the regressors in all time periods in the equation, in the spirit of Chamberlain (1984). This substantially complicates the analysis, however, particularly due to the unbalanced nature of the panel. In specifications that include log financial wealth as a regressor, we will also include the sample average of log wealth. The robustness of our results to correlations between α_1 and α_2 and other regressors will be investigated using a linear probability model, discussed at the end of the next section.

Initial Conditions and Estimation

This subsection is an informal discussion of how to estimate the model; details can be found in Alessie *et al.* (2001). In a short panel, there is a problem with the initial conditions (cf. Heckman (1981a)). One way to deal with this problem is to add static ("reduced form") equations for the first time period similar to the dynamic equations, but without the lagged dependent variables. The coefficients are allowed to be different from the coefficients in the dynamic equations, the random effects are linear combinations of the random effects in the dynamic equations, and the error terms are allowed to have a different covariance structure. This is the straightforward generalization of the solution that was given by Heckman (1981b) for the univariate case. In principle, the static equations can be seen as linearized approximations of the true reduced form (obtained by recursively eliminating y_{t-1} until $t = -\infty$). Heckman's simulations suggest that the procedure already works well in short panels, i.e. the approximation error does not lead to a large bias on the parameter estimates.[3]

The complete model can then be estimated by Maximum Likelihood (ML), including the nuisance parameters of the static equations. Conditional on the random effects, the likelihood contribution of a given household can be written as a product of bivariate normal probabilities for over all time periods. Each bivariate normal probability is then the probability of the observed ownership state, either conditional on the ownership state in the previous year ($t \geq 2$) or unconditional ($t = 1$).

Since random effects are unobserved, the actual likelihood contribution is the expected value of the conditional likelihood contribution, with the expected value taken over the two individual effects. This is a two-dimensional integral. It can be approximated numerically using, for example, Gauss–Hermite–quadrature. Instead, we use simulated ML: bivariate errors are drawn from $N(0, I_2)$, they are transformed into draws of the random effects using

[3] An alternative solution is explored by Lee (1998), who treats the initial values as fixed. Lee's simulation evidence suggests that this does not lead to any serious bias if the panel consists of 20 waves, but it does if the panel has only eight waves. It therefore seems less appropriate for our panel of six waves. Chay and Hyslop (2000) compare various ways to deal with the initial conditions problem in logit and probit models. They find that the probit model with the Heckman procedure performs better than other random effects models.

the parameters of the random effects distribution, the conditional likelihood contribution is computed for each draw, and the mean across R independent draws is computed. If $R \to \infty$ with the number of observations, this gives a consistent estimator; if draws are independent across households and $R \to \infty$ faster than $\sqrt{N}$, then the estimator is asymptotically equivalent to exact ML (see Hajivassiliou and Ruud (1994), for example).[4]

In our case, the data at hand is an unbalanced panel, due to attrition, non–response, and refreshment. We assume that attrition and item non–response are random. We will use the complete unbalanced panel. This is more efficient than using the balanced sub–panel only.

3 Data

We use six waves of the CentER Savings Survey (CSS), drawn from 1993 until 1998. Nyhus (1996) describes the set up of this data set and its general quality. The panel consists of two samples. The first is designed to be representative of the Dutch population (REP), but, due to survey non-response, the actual REP samples are not completely representative. The REP contains approximately 2000 households in each wave, including refreshment samples compensating for panel attrition. The second sample was drawn from high-income areas and should represent the upper income decile (HIP). Initially, it consisted of about 900 families. It is available in each wave except the final one. For our analysis, REP and HIP samples are combined. In the descriptive statistics, we correct for non-random sampling by using sample weights that are based upon income and home ownership. These weights are constructed using information from a much larger data set (Housing Needs Survey (WBO), collected by Statistics Netherlands), which is close to representative for the Dutch population.

The CSS data were collected via on-line terminal sessions, where each family was provided with a PC and modem. The answers to the survey questions provide general information on the household and its members, including work histories and labor market

[4]In the application, we found $R = 100$ to be sufficiently large in the sense that results did not change if R was increased further.

status, health status, and many types of income. Important for our purposes are the questions on assets and debts. For most of the forty asset and debt categories, respondents first indicate whether they own the type. If they do, they get a series of questions on amounts and the precise nature of each asset in that category. Non–response in the ownership questions is negligible, but non–response in some of the questions on the amounts is substantial. On average, about 20% of those who own stocks do not know or refuse to give the value of their stocks. Mutual funds have a lower non–response rate of around 13% per year. For some descriptive statistics (such as shares of specific asset types in total financial assets, see below), the item non-response creates a problem. We have therefore imputed the amounts for those who reported to be owners but did not provide an amount. See Alessie *et al.* (2002), who also provide an extensive description of all categories of assets and debts in the survey.

We focus on two types of risky financial assets: stocks and mutual funds. The CSS distinguishes between two types of stocks: stocks from substantial holding and (other) shares of private companies. There are very few people who hold the former type, but these people typically hold high amounts. The two types of stocks are different for tax purposes, since income from a substantial holding is treated as business capital. There is a large variety of mutual funds, composed of many different assets. In 1997, according to De Nederlandsche Bank (2001), more than 50% of the total amount in mutual funds was invested in stocks, while about 30% was held in real estate and about 10% in bonds. The survey data that we use do not provide enough information to separate the mutual funds according to their investment.[5]

During the period under consideration, dividends from other shares and from mutual funds were liable to income tax to the extent that they exceed an exemption threshold (Dfl 2,000 for couples, Dfl 1,000 for singles). Capital gains on these are not taxed. The thresholds on dividends are separated from the thresholds on interest on savings, creating a tax incentive for holding stocks or mutual funds as well as saving accounts.

[5]Note that the definition of mutual funds deviates from that adopted in Alessie *et al.* (2002) where so–called ‘growth funds’ are included. The latter are typically low–risk funds, which is the reason why we exclude them here.

The first two columns of Table 1 show how ownership rates of the two types of assets have developed during the years of the survey. The ownership rate of stocks has risen from 11.4% to 15.4%. Mutual funds were more often held than stocks, with an even higher growth rate during the sample period. Many financial institutions have been successful in presenting mutual funds as a low threshold asset, available to many individual investors. Still, the majority of Dutch households held neither stocks nor mutual funds in 1998. This lack of participation can be explained by monetary transaction costs and information costs, both of which can be substantial.[6]

The remaining columns of Table 1 show the time path of amounts invested in stocks and mutual funds, as shares of total financial assets.[7] While the ownership rate of stocks is always lower than the ownership rate of mutual funds, the reverse is true for the shares of stocks and mutual funds in total financial wealth. This is because the few people who hold stocks typically hold high amounts of them. The growth of the shares is less spectacular than the growth of the ownership rates. The shares may be strongly influenced by some large amounts, due to the skewed distribution of wealth and its components. Some rich people hold large amounts, and there are very few of these in the sample, particularly in 1998, the year without high income panel. This may explain why some of the time patterns are not as pronounced as in aggregate data produced by Statistics Netherlands (see Alessie *et al.* (2002)). In the remainder of this study, we will not use the amounts data and focus on ownership.

In Figures 1 and 2, we present (head of household) age and cohort patterns of the ownership rates of stocks and mutual funds, based upon the six waves of the survey. We use five year–of–birth cohorts, with birth years 1915–1919 for the oldest cohort, until birth

---

[6] In the Netherlands, explicit transaction costs are low (about 0.5% of the investment) but implicit costs (entry and exit fees incorporated in the buying and selling price of the fund) are higher. The maximum entry fee is about 2.5% of the investment, and the maximum exit fee is about 1.5% (see Consumentenbond (1999)). Apart from the transaction costs, most mutual funds charge a management fee of about 0.5% per year and apply minimum investment restrictions. These implicit costs are comparable to the substantial transaction costs in Italy discussed by Guiso and Jappelli (2002). It is not clear, however, whether Dutch investors are aware of the implicit costs.

[7] The share is defined as the total amount invested in the asset by all households (weighted with the sample weights), divided by (weighted) total financial wealth of all households.

years 1970–1974 for the youngest cohort. Cohort labels indicate the middle year-of-birth. Each figure gives the raw ownership rates for each cohort in each wave; the six points for each cohort represent the six average age levels at the times of the six interviews, and form a "cohort curve". The jumps between the cohort curves show that, apart from age effects, there are cohort or time effects. The cohort curves are not horizontal, implying that there are time and/or age effects; the fact that not all cohort curves are the same shows that there is more than just time effects. As usual, cohort, time and age effects cannot be identified without further assumptions. A plausible interpretation of both figures, assuming that cohort effects are zero, is that ownership rates increase with age and that there are positive effects of calendar time, particularly for the older cohorts. Alessie *et al.* (2002) find a similar increasing age pattern for the category of all risky financial assets. This deviates from the pattern for some other countries. Italy and the US, for example, have a hump shaped pattern.

Table 2 describes the dynamics of the ownership patterns of stocks and mutual funds separately. This gives a partial view of mobility, since only transitions of households that sell all their stocks or mutual funds or enter the market of stocks or mutual funds are shown, and not the changes in (positive) amounts held. The table presents the numbers of owners and non-owners of stocks and mutual funds in each balanced subsample and the percentages of exits and entries.[8] For example, 4.2% of households that did not own stocks in 1993, owned stocks in 1994. This is about 3.6% of all households in the 1993–1994 sample. On the other hand, 22.6% of those who owned stocks in 1993 (i.e., 3.4% of all households in the sample) no longer owned stocks in 1994. Thus ownership mobility is substantial, for stocks as well as mutual funds. In particular, the fractions of owners selling their mutual funds or stocks are larger than expected, given the high returns on these assets in the nineties. On average, 21.2% of all stock owners no longer owned stocks one year later, and 26.3% of mutual fund owners no longer owned mutual funds one year later. The large transition rates are in a similar order of magnitude as those reported for the US and Italy.[9] In absolute numbers or as a percentage of total sample size, entry

[8]The sampling weights are not used. The transition rates hardly change if observations with very small amounts are excluded.

[9]Vissing–Jørgensen (2002, p. 13) finds that in the Panel Study of Income Dynamics, 28.1% of all

transitions are somewhat more numerous than exit transitions, in line with the rising ownership rates in Table 1. Entry and exit numbers are larger for mutual funds than for individual stocks, which could point at lower transaction costs for mutual funds.

The balanced subpanel of households that participated in the survey for six consecutive years consists of 405 households. 297 of these never owned stocks, while 34 households owned stocks in all six waves. 271 households never owned mutual funds and 21 always owned mutual funds. In this subsample, there are 47 households who exit from stocks at least once, and 23 of these go back into stocks after they have exited. 76 households exit from mutual funds at least once, and 35 of them go back into mutual funds after having exited.

Where did people go after exiting from stocks or mutual funds? Did stockholders leave the market completely or did they become mutual fund holders? Averaging over all pairs of waves shows that of those who exited from stocks and did not own mutual funds, only 12.1% went into mutual funds. Of those who exited from mutual funds and did not own stocks, 14.1% went into stocks. Complete substitution of stocks for mutual funds or vice versa is thus not very common; most people who sell off their only type of risky financial assets leave the market completely.

Table 3 shows the correlation between holding one asset type in one period, and holding the other asset type in the next period. For all years, the ownership rate of stocks in year $t+1$ is larger for those with mutual funds in year t than for those without mutual funds in year t – conditional on not owning stocks in year t. For example, 9.4% of those without stocks and with mutual funds in 1993 owned stocks in 1994. On the other hand, only 3.5% of those who had neither stocks nor mutual funds in 1993, owned stocks in 1994. Thus there is some positive correlation across ownership of the two asset types. The same conclusion is obtained when ownership rates of mutual funds are considered. Whether

households hold stocks in 1989 but not in 1994 or vice versa. Kennickell and Starr-McCluer (1997, p. 455) consider ownership of one category consisting of stocks, mutual funds, managed investment accounts or trusts in the Survey of Consumer Finances 1983–1989, and report transition rates of 10% from ownership to non–ownership and 19% from non–ownership to ownership. Miniaci and Ruberti (2001) report two-years transition rates from ownership to non-ownership between 32% and 42% using the SHIW survey of the Bank of Italy.

this positive correlation reflects some genuine state dependence effect (such as learning) or (observed or unobserved) heterogeneity, is one of the issues we will analyze in the next section, using the models in Section 2.

4 Results

The results of the random effects probit model are discussed in the first subsection. In the second subsection, these results are briefly compared to those of a fixed effects linear probability model. In the final subsection, the implications of the probit results for explaining the growth in ownership rates of stocks and mutual funds are presented. This is based on predicted probabilities, which are not always between 0 and 1 in the linear probability models, and will therefore be done on the basis of the probit results only.

4.1 Random Effects Probit

Tables 4 and 5 give the results for four specifications of the bivariate probit model. For stocks and mutual funds, the same explanatory variables are used. Financial wealth is not included in the first two specifications since it may not be strictly exogenous. In the final two specifications, lagged log financial wealth and its own household specific average (over the observation window) are included. This controls for the potential correlation between lagged financial wealth and the individual effects (see Hausman and Taylor (1981), for example).[10] We find a very strong positive relation between the average log financial wealth level on ownership of stocks as well as mutual funds. The effect of log financial wealth in the previous year is much smaller and, surprisingly, negative. It is significant at the 5% level for stocks and at the 10% level for mutual funds. Probably the relation between asset ownership and financial wealth requires more structure than is incorporated in the current model, and this is beyond the scope of the paper. The main reason for presenting results with financial wealth is to check whether or not estimates are sensitive to its inclusion.

[10]Including an arbitrary linear combination as in Hyslop (1999) is not possible due to the unbalanced nature of the panel.

In specifications 2 and 4, an indicator for risk aversion is included. This is the answer on a scale from 1 (*strongly disagree*) to 7 (*agree strongly*) to the question whether the head of household agrees with the statement *I think it is more important to have safe investments and guaranteed returns than to take a risk to have a chance to get the highest possible returns.* This question has only been posed to respondents with household income of 20,000 Dfl. or more, thus excluding 15.6% of all households. Moreover, 4.8% of all households in the sample answered "don't know". As expected, the probability to own stocks falls with the degree of risk aversion. Such a relation is not found, however, for mutual funds. In specification 2, those who answer "don't know" or do not get the question have a lower tendency to own mutual funds. In specification 4 where financial wealth is also controlled for, the risk aversion dummies are jointly insignificant in the mutual funds ownership equation.

To avoid correlation with random effects and endogeneity of income and the marginal tax rate, income is non-capital income and the marginal tax rate is the maximum of the within–household imputed marginal rate applied to pseudo–taxable income, in which individual capital income is replaced by its cross–sectional average (following Agell and Edin (1990)). The effects of income and the marginal tax rate are hard to disentangle, due to the strong (positive) correlation between these variables. We find that both effects are positive for both types of assets. For stocks, the income effect is significant (at the two-sided 5% level), while for mutual funds, the tax effect is significant. An explanation for the stronger income effect for stocks than for mutual funds may be that high income households will typically have more to invest, making the relatively large fixed costs component of buying or holding individual stocks less important. Moreover, the sheer size of the portfolios of the higher income groups may be bigger relative to the unit prices of stocks, thus allowing them to hold a bigger variety of stocks and to achieve greater diversification than low income groups. This argument is not present in mutual funds where holders can achieve diversification with small investment amounts. King and Leape (1998) find that the marginal tax rate has a statistically significantly positive effect on ownership of corporate equity but conclude that taxes play a rather limited role in explaining differences in portfolio composition across households, especially when it comes to stock ownership.

The income tax rules for stocks and mutual funds are the same (see Section 3). The fact that capital gains are not taxed creates an incentive to hold stocks or mutual funds, which increases with the household's marginal tax rate.[11] This explains the positive effect of the marginal tax rate. The larger tax effect for mutual funds could be due to the fact that suppliers of these funds strongly advertise their tax favored nature.

Labor market status variables for the head of household are jointly significant in both equations. The retired are significantly more likely to own stocks or mutual funds than employees. The reason might be that income risk from pensions is lower than from earnings, and that the willingness to invest in risky assets falls with the amount of background risk (cf., e.g., Hochguertel (2003)).

The most striking result is the enormous effect of self–employment on ownership of stocks: a self–employed head has a more than 25%-points higher probability to own stocks than an employee (the reference group), *ceteris paribus*, which is at odds with findings in other countries.[12] The effect remains almost the same if controls for financial wealth or risk aversion are included. This result is consistent with Heaton and Lucas (2000) who find that people with private business holdings in the US hold a disproportionately large share of common stocks. Part of the explanation could be that the self–employed often hold shares in their own firm which will often be shares from a substantial holding. Excluding stocks from a substantial holding from our analysis reduces the size of the effect to about half its size, though it remains strong and significant.

Since Figures 1 and 2 in the previous section have a plausible interpretation without cohort effects, we have included age and time effects but no cohort effects. This identifies the age and time patterns. Age is significantly positive for stocks as well as mutual funds, in line with findings for risky asset ownership by King and Leape (1987), who attribute the age effect to the accumulation of information about investment opportunities. This information argument seems particularly relevant for individual stocks, since these are the more "information intensive" type of risky assets. The time effects are similar for the

[11]See Poterba (2002) for a general discussion of the impact of tax rules on portfolio choice.

[12]Bertaut and Starr-McCluer (2002) and Eymann and Boersch-Supan (2002) find that the self-employed are less likely to hold risky financial assets than employees in the US and in Germany.

two asset types and show that the assets have become more popular during the last few years of the survey.

The education coefficients indicate that the higher educated significantly more often hold stocks but not mutual funds. Again, this could be attributed to the greater information processing requirements for investing in stocks relative to mutual funds where investments are made by professionals. If financial wealth is included, however, the effects of education vanish (see Table 5), implying that the education effects in Table 4 might pick up wealth effects. This finding is quite unusual in the literature (cf., e.g., Guiso and Jappelli, 2002, and Bertaut and Starr-McCluer, 2002). A similar interpretation can be given for the dummy "High Income Panel." The positive significant effect of this dummy for both asset types largely vanishes if financial wealth is included (Table 5). The way the high income sample is drawn and the general tendency of the income–rich to save a lot make it plausible that selection into this panel is not only based upon income but also on wealth, explaining why the dummy variable serves as a wealth proxy.

The estimated standard deviations of the random effects are 1.44 and 1.20 for stocks and mutual funds, respectively. The standard deviations of the error terms are normalized to one. Thus, unobserved heterogeneity plays a major role, explaining more than half of the unsystematic variation in the model.

In both equations, the lagged dependent variables concerning ownership of the same asset type are significantly positive. To interpret these results, predicted ownership probabilities for the various lagged ownership states are presented in the final panels of Tables 4 and 5. Exogenous variables are set to their (weighted) sample means and random effects are set to zero. According to specification 1, owners of stocks are about 16.7% points more likely to own stocks next period than non–owners with the same (observed and unobserved) characteristics if they do not own mutual funds, and 15.4% points if they do. For mutual funds, the differences are even larger (20.1% points if no stocks are held, 17.4% points if stocks are held). Similar differences are found for the other specifications. Explanations for positive state dependence are the costs of acquiring stocks or mutual funds (i.e., genuine transaction costs, not the costs of holding the assets)[13] and asset spe-

[13]Hyslop (1999) formalizes this in a stylized dynamic optimization model; a similar model can be used

cific learning: once people own the asset, they are more familiar with it, and are more aware of its risk and return characteristics.

All the results discussed so far relate to the dynamics of each of the two types of assets separately. In most respects, these results are similar to what would be predicted by separate univariate models. The bivariate model, however, also gives insight in the relation between the two ownership decisions.

The "cross–effects" of lagged ownership of one asset type on ownership of the other asset type are both negative and one of them is significant at the 5% level: *ceteris paribus*, those who do not own stocks are significantly more likely to own mutual funds in the next period than those who own stocks. According to specification 1 in Table 4, the differences are 4.4%-points and 2.9%-points for those who did and did not own mutual funds in the previous period. If financial wealth is controlled for (specification 3 in Table 5), the other cross–effect becomes significantly negative also. If the risk aversion index is included (specifications 2 and 4), the negative signs remain but the effects are no longer significant.

The negative cross–effects cannot be explained by general learning: if ownership of one asset type would improve knowledge about the other asset type, a positive cross–effect would result. Thus the results on the own and cross–effects together suggest that if learning takes place, it is specific for the type of asset.

On the other hand, negative cross–effects are also consistent with the adjustment cost arguments that explained the strong positive effects of lagged ownership of the same asset type. People who own stocks but no mutual funds have an incentive to remain focused on stocks to avoid the adjustment costs, while people who own neither stocks nor mutual funds and who consider investing in risky assets, face adjustment costs anyhow. Adjustment costs thus give an explanation for own as well as cross state dependence effects, while learning can only explain the univariate effects. These adjustment costs may reflect the actual (monetary) transaction costs involved with buying or selling an asset, but may also include non-monetary components such as the required effort, the need to collect information, etc. Sticking to one type of assets may also be due to the realization that too much churning does not pay and that portfolio choice should be

here for each of the two assets separately.

governed by a long horizon.

The estimated correlation coefficient between the two random effects is significantly positive, ranging from 0.42 in specification 4 to 0.66 in specification 1. This suggests that the people who have a large preference for holding stocks (given their observed characteristics), tend to be the same people who have a preference for holding mutual funds. These may be the people with lower degrees of risk aversion or higher interest in financial markets. The positive correlation between holding stocks and holding mutual funds in the data, is to a large extent due to this positive correlation in unobserved heterogeneity.

Allowing for correlation in the individual effects in the two equations has a major impact on the estimates of the cross–effects of ownership of one type of asset on ownership of the other type of asset in the next time period. If we estimate the model with the correlation between the random effects restricted to zero, we find significant positive estimates for both cross–effects. Not allowing for correlation between unobserved heterogeneity terms would thus lead to a large upwards bias on the effect of ownership of one asset type on ownership of the other type and to misleading conclusions about the importance of general learning.

The correlation between the error terms in the two equations is small and insignificant. A negative correlation could point at fixed holding costs for each asset type (such as monitoring costs) that would be an incentive for specialization. Vissing–Jørgensen (2002) finds evidence of such costs. A positive correlation could point at a common element in monitoring both assets, or at benefits of diversification. Apparently, the positive and the negative effects cancel or do not play a role.

4.2 Linear Probability Model

To check the sensitivity of the main conclusions for the random effects assumption, some standard linear dynamic panel data models were estimated. See, for example, Verbeek (2000, Section 10.4) for an accessible exposition and Hyslop (1999) for a comparison of univariate linear probability and random effects models. The following assumptions are the basis for GMM estimation:

1. The time varying independent variables are uncorrelated to the error terms (strict exogeneity) but can be correlated with the individual effects.

2. Time invariant independent variables included in the model are uncorrelated to the error terms as well as the individual effects.[14] [15]

3. The error terms are uncorrelated over time.

To avoid the well–known problem that the small sample performance of GMM deteriorates with too many moments, only the moments combining regressors and error terms that are as close as possible in time are used. Standard GMM estimation is applied, separately for the equations for stocks and mutual funds. Any type of heteroskedasticity is allowed for, including that implied by the binary nature of the dependent variable.

We ran several specifications with different sets of independent variables.[16] On the basis of Sargan tests for the over–identifying restrictions, we selected the specification presented in Table 6. For stocks, the over–identifying restrictions are not rejected at the 2% level although they are rejected at the 5% level. For mutual funds, the overidentifying restrictions are not rejected at any conventional significance level. Moreover, the hypothesis of no second order autocorrelation in the residuals of the differenced equations is not rejected for either type of assets, supporting the assumption of no autocorrelation in the error terms.

The same explanatory variables are used as in the probit model in Tables 4 and 5. The results are largely in line with the probit results. For example, the large effect of self–employment on the probability of holding stocks is again the most salient finding. We can now conclude that this is not an individual effect, since correlation between individual effects and regressors is controlled for.

The main difference with specification 1 in Table 4 is the estimated "cross–effect" of lagged ownership of stocks on ownership of mutual funds, which is now positive and insignificant, whereas it was negative and significant in the random effects probit model.

[14]This also applies to some variables that only vary systematically over time such as age.

[15]We have also estimated a model in which this assumption is avoided, excluding the time-invariant variables from the regression. This exclusion has little effect on the other parameter estimates.

[16]We use the DPD98 software as described in Arellano and Bond (1998).

This confirms that no evidence of general learning is found, but does not support the adjustment costs argument given in the previous subsection.

The main purpose of the linear probability models is to perform a sensitivity check on the findings on the basis of the probit models. The specification in Table 6 as well as estimates of alternative specifications confirm that most findings are very similar (tax and income effects, effect of labor market position, effects of lagged ownership of an asset type on the same asset type). The only differences concern the cross–effects of lagged ownership of stocks on ownership of mutual funds and of lagged ownership of mutual funds on ownership of stocks. Still, we always find significant negative effects or insignificant effects, and only find significant positive effects if a zero correlation of unobserved heterogeneity terms for the two equations is imposed—a restriction that is always fiercely rejected in favor of positive correlation. This confirms that cross–effects due to general learning cannot be established and points at asset specific learning and adjustment costs.

4.3 Explaining the Growth in Ownership Rates

The probit model results presented in Table 4 can be used to predict ownership probabilities for individual households and aggregate ownership rates for groups of households under different scenarios. Such predictions can be used to analyze how much the explanatory variables in the equations contribute to the changes in the aggregate ownership rates of stocks and mutual funds over time. The idea is similar to the Oaxaca decomposition that is commonly used in studies on wage differentials (Oaxaca, 1973).

The results for specification 1 are presented in Table 7. The top panel refers to stocks, the bottom panel to mutual funds. The first row of each panel presents observed changes in mean predicted ownership rates (in percentage points), using common samples for the two years considered. Random effects are set to zero. Since time dummies for all years are included, these sample changes reflect the increasing trend in aggregate ownership rates over the same years reasonably well.[17] The other rows compare two sets of predicted

[17]This should still improve if random effects were integrated out instead of set to zero, but then the correlation between random effects and lagged ownership dummies should be accounted for.

aggregate ownership rates: those using the observed explanatory variables and those in which one or more explanatory variables are replaced by their lags. Take, for example, the change in the stocks ownership rate from 1997 to 1998 of about 3.48% points. If in the 1998 sample age is replaced by its lagged value (i.e., age in 1997), the predicted mean ownership rate falls by 0.26% points. In other words, age changes explain 0.26% points of the total 3.48% points. Similarly, changes in incomes and marginal tax rates (which become somewhat larger, on average) explain a 0.20% points rise of the ownership rate. All exogenous regressors (not including time dummies or lagged dependent variables) explain a rise of about 0.44% points. The lagged dependent variables explain a rise of 0.11% points, mainly due to the rise in the ownership rate of stocks from 1996 to 1997. In total the regressors in the model (time dummies not included) explain 0.55% points of the 3.48% points rise in ownership of stocks. The remainder is not explained by the regressors and captured by the time dummies, whose contribution can be seen as the (residual) part of the change in ownership that cannot be explained by the economic variables in the model.[18]

The results for mutual funds are similar. The conclusion is that age is the only exogenous variable which consistently positively contributes to explaining the rising ownership rates. The main reason is that age is not only significant in the probits but also systematically increases over time. Self-employment, for example, is very important for ownership of stocks, but the fraction of self–employed in the sample does not vary systematically over the years.

For the other specifications, the regressors explain a similarly small part of the total rise in the ownership rates. According to specification 4, the fall in average risk aversion explains 0.29% points and 0.48% points of the rise in ownership of stocks in 1996–1997 and 1997–1998, but still, all the economic variables together explain only a small part of the total rise. This may mean that most of the increase was induced by supply factors such as the development of the mutual fund industry, aggressive advertising of financial services, etc., rather than by changes in household circumstances.

[18]This can be compared to the changes "explained by" parameter changes in the usual Oaxaca decomposition; in this model with time dummies, only the constant term can vary over time.

5 Conclusions

As the stockholder base has widened considerably over the last decade in many countries, understanding how households make their portfolio decisions over time has wide–ranging implications for understanding the allocation of risk in financial markets, the distribution of wealth, and pricing relationships for individual assets. This paper is one of the first studies of the dynamics of individual households' (multivariate) investment strategies using representative panel survey data.

We have estimated dynamic models explaining ownership of the two main types of risky financial assets in the Netherlands: stocks and mutual funds, that differ in degree of riskiness, "information intensiveness," and adjustment costs. Mutual funds are particularly attractive for small investors with little financial knowledge, since they are managed by professionals. Our results confirm this to some extent, since we find that the probability to own stocks increases significantly with income while the probability to own mutual funds does not. Tax incentives, on the other hand, play a larger role for mutual funds than for stocks. Self–employed are much more likely to hold stocks than others, while they do not have a different ownership rate of mutual funds. Explanations could be that the self–employed are interested in specific stocks to hedge against their larger income uncertainty, or that acquiring knowledge about specific firms and their prospects is closer to their every day practice. The alternative explanation that the self–employed simply have different preferences and care less about diversification, is unlikely since the effect remains the same if a subjective index for risk aversion is included or if unobserved (preference) heterogeneity is allowed to be correlated with background variables in a fixed effects setting.

We find that the dynamics of ownership of either type of risky assets are driven by state dependence as well as unobserved heterogeneity. Both explain part of the persistence of ownership of both types of assets in the data. The state dependence can be explained from the adjustment costs of buying or selling the asset or from asset type specific learning. On the other hand, the positive sample correlation between ownership of one type of asset and lagged ownership of the other asset is explained from (observed and unobserved) heterogeneity only. One source of this could be a joint element in monitoring or other holding costs that makes it attractive to hold both asset types simultaneously. Another reason could be that combining stocks and mutual funds creates opportunities for diversification that cannot be attained by mutual funds alone (since these typically invest in certain sub-samples of stocks).

If correlation between the unobserved heterogeneity in the two ownership equations is not allowed, a positive effect of owning one type of asset on owning the other type in the next year is found. This changes completely, however, once the two unobserved heterogeneity terms are allowed to be correlated. We then find a positive correlation and no evidence that households substitute one type of assets by the other, or that ownership of one type leads to general learning, more financial knowledge and a larger probability of buying the other type of assets. In contrast, we find some evidence of a negative effect of owning one type of assets on buying or keeping the other type. This can be explained by adjustment costs, implying that those who have acquired one specific asset will tend not to reallocate their money to the other type of assets. Such adjustment costs will comprise the actual transaction costs involved with portfolio adjustment, but may also contain non–monetary or perceived costs components, reflecting effort, the cost of acquiring information, etc..

In their introduction Guiso *et al.* (2002) report that, conditional upon ownership, the age profile is relatively flat in most countries. This finding indicates that people do not rebalance their wealth portfolio very frequently. Guiso *et al.* (2002) point out that this portfolio inertia cannot be justified by (standard) analytical or computational portfolio models. Analytical models stress the role of the length of horizon for optimal choice of portfolio composition (see Gollier (2002)). More complicated simulation models similarly imply policy functions for portfolio shares that are sensitive to age (see Haliassos and Michaelides (2002)). Moreover, professional financial advisors encourage their clients to reduce the share of risky assets as they age. Ameriks and Zeldes (2001) also find that households choose a particular portfolio of assets and do not make many changes even when their circumstances change and in spite of professional advice. Similarly, Poterba (2002) finds that households fail to rebalance their portfolios to take full advantage of (changes in) tax provisions.

Classical papers like those of Merton (1969) and Samuelson (1969) assume that agents live in a frictionless world and have HARA preferences. They predict myopic optimal behavior (i.e. a constant fraction of risky assets in the portfolio) for a given individual. To make such models more realistic, it would be useful to introduce dynamic features. Our results would suggest asset specific learning experiences and real or perceived adjustment costs are particularly relevant. These can explain why households once they have decided which asset classes to participate in, are not likely to revisit this decision.

Future research can go in several directions. First, we have not modelled the amounts held. Although this is not without measurement problems, it certainly seems a relevant extension. It could also help analyze the importance of fixed costs of holding, buying, and selling assets, extending the work of Vissing–Jørgensen (2002). Second, if data for a longer time period become available, it seems useful to relate the ownership dynamics to the trends in the financial markets or to relevant macro variables such as unemployment, inflation or expected inflation, consumer confidence, etc. Third, straightforward extensions of our models could be used to analyze other asset and debt types, or to analyze assets at a less aggregate level. For example, to understand the dynamics and in particular the underlying cost structure driving these, it seems relevant to distinguish people who substitute one stock for the other from people who do not trade at all.

Finally, it would be interesting to extend the analysis with data on the recent time period with falling asset returns. The rationale for people then not venturing into risky assets seems to be that the return foregone is too low compared to the costs saved. In times when returns are high (1993–1998), this may make sense, and the time dummies in our models seem to pick up the sluggishness in the adjustment. On the other hand, there is some asymmetry in the sense that once adjustment costs are incurred, part of them will be sunk (at least the information acquisition costs). Selling the assets will therefore not be associated with the same costs. Such an analysis will have to await future data releases covering the recent period of downturn.

References

Agell, Jonas and Per–Anders Edin, "Marginal Taxes and the Asset Portfolios of Swedish Households," *Scandinavian Journal of Economics,* 92:1 (1990), 47–64.

Alessie, Rob, Stefan Hochguertel and Arthur van Soest, "Ownership of Stocks and Mutual Funds: A Panel Data Analysis," Discussion Paper #2001–94, CentER for Economic Research, Tilburg University (2001).

Alessie, Rob, Stefan Hochguertel and Arthur van Soest, "Household Portfolios in the Netherlands", in Guiso *et al.* (2002), 341–388.

Ameriks, John and Stephen P. Zeldes, "How Do Household Portfolio Shares Vary with Age?" working paper, Columbia University, New York (2001).

Arellano, Manuel and Stephen R. Bond, "Dynamic Panel Data Using DPD98 for GAUSS," mimeograph, Institute for Fiscal Studies (1998).

Bertaut,Carol C. and Martha Starr-McCluer, "Household Portfolios in the United States," in Guiso *et al.* (2002), 181–217.

Chamberlain, Gary, "Panel Data," in Zvi Griliches and Michael D. Intriligator (eds.), *Handbook of Econometrics*, (Amsterdam: North-Holland, 1984), 1248–1318.

Chay, Kenneth Y. and Dean R. Hyslop, "Identification and Estimation of Dynamic Binary Response Panel Data Models: Empirical Evidence using Alternative Approaches," mimeograph, University of California at Berkeley (2000).

Consumentenbond, *Jaarboek Beleggen 1999*, (Utrecht: Kosmos, 1999).

De Nederlandsche Bank, "Recordinleg bij beleggingsinstellingen in 2000," *Statistical Bulletin Central Bank of the Netherlands* (2001), 17–22.

Eymann, Angelika and Axel Boersch-Supan, "Household Portfolios in Germany," in Guiso *et al.* (2002), 291–340.

Guiso, Luigi, and Tullio Jappelli, "Household Portfolios in Italy", in Guiso *et al.* (2002), 251–289.

Guiso, Luigi, Michael Haliassos and Tullio Jappelli, eds., *Household Portfolios* (Cambridge (Mass.): MIT Press, 2002).

Hajivassiliou, Vasilis A. and Paul A. Ruud, "Classical Estimation Methods for LDV Models Using Simulation," in: Robert F. Engle and Daniel L. McFadden (eds.), *Handbook of Econometrics, vol. IV* (New York: North–Holland, 1994), 2384–2443.

Haliassos, Michael, and Alexander Michaelides, "Calibration and Computation of Household Portfolio Models," in Guiso *et al.* (2002), 55-102.

Hausman, Jerry A. and William E. Taylor, "Panel Data and Unobservable Individual Effects," *Econometrica*, 49:6 (1981), 1377–1398.

Heaton, John, and Deborah Lucas, "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk," *Journal of Finance*, 55:3 (2000), 1163–1198.

Heckman, James J., "Statistical Models for Discrete Panel Data," in Charles F. Manski and Daniel L. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications* (London: MIT Press, 1981a), 114–178.

Heckman, James J., "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in Charles F. Manski and Daniel L. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications* (London: MIT Press, 1981), 179–195.

Hochguertel, Stefan, "Precautionary Motives and Portfolio Decisions," *Journal of Applied Econometrics*, forthcoming (2003).

Hyslop, Dean R., "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women," *Econometrica*, 67:6 (1999), 1255–1294.

Ioannides, Yannis M., "Dynamics of the Composition of Household Asset Portfolios and the Life Cycle," *Applied Financial Economics*, 2:3 (1992), 145–159.

Kennickell, Arthur B. and Martha Starr-McCluer, "Retrospective Reporting of Household Wealth: Evidence from the 1983-1989 Survey of Consumer Finances," *Journal of Business & Economic Statistics*, 15:4 (1997), 452-463.

King, Mervyn A. and Jonathan I. Leape, "Asset Accumulation, Information, and the Life Cycle," NBER Working Paper, No. 2392 (1987).

King, Mervyn A. and Jonathan I. Leape, "Wealth and Portfolio Composition: Theory and Evidence," *Journal of Public Economics*, 69:2 (1998) 155–193.

Lee, Lung–Fei, "Simulated Maximum Likelihood Estimation of Dynamic Discrete Choice Statistical Models Using some Monte Carlo Results," *Journal of Econometrics*, 82:1 (1998), 1-35.

Merton, Robert C., "Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case," *Review of Economics and Statistics*, 51:3 (1969), 247–257.

Miniaci, Raffaele and Francesca Ruberti, "Household Risky Asset Ownership: A Dynamic Analysis," mimeograph, Università di Padova 2001).

Nyhus, Ellen K., "The VSB-CentER Savings Project: Data Collection Methods, Questionnaires and Sampling Procedures," VSB-CentER Savings Project Progress Report no. 42, Tilburg University (1996).

Oaxaca, Ronald L., "Male–Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14:3 (1973), 693-709.

Poterba, James M., "Taxation and Portfolio Structure: Issues and Implications," in Guiso *et al.* (2002), 102–142.

Samuelson, Paul A., "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51:3 (1969), 239–246.

Verbeek, Marno, *A Guide to Modern Econometrics*, (Chichester: Wiley, 2000).

Vissing–Jørgensen, A., "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures," NBER Working paper W8884, National Bureau of Economic Research, Cambridge (Mass.) (2002).

Tables and Figures

Table 1: Ownership Rates and Portfolio Shares

	ownership rate		portfolio share	
year	stocks	mutual funds	stocks	mutual funds
1993	11.4	11.8	21.3	5.4
1994	9.9	12.8	20.6	6.7
1995	11.4	12.9	22.0	6.2
1996	13.5	14.7	24.0	7.0
1997	14.4	16.2	25.3	7.1
1998	15.4	18.4	23.8	10.0

Note: weighted statistics; portfolio share: ratio of wealth held in stocks or mutual funds to total financial assets

Table 2: Transitions (Univariate)

			Stocks			
years $t/t+s$	owners at t	exits in % of owners	exits in % of all	non-owners at t	entries in % of non-owners	entries in % of all
1993/94	310	22.6	3.4	1768	4.2	3.6
1994/95	290	26.6	4.1	1593	5.5	4.7
1995/96	275	21.5	3.5	1407	6.1	5.1
1996/97	239	19.7	3.3	1174	6.0	5.0
1997/98	159	15.7	2.5	839	4.4	3.7
1993/97	128	25.8	4.1	672	11.6	9.8
			Mutual Funds			
years $t/t+s$	owners at t	exits in % of owners	exits in % of all	non-owners at t	entries in % of non-owners	entries in % of all
1993/94	310	26.5	4.0	1768	6.8	5.8
1994/95	336	32.4	5.8	1547	6.1	5.1
1995/96	299	20.4	3.6	1383	5.7	4.7
1996/97	253	26.1	4.7	1160	8.2	6.7
1997/98	190	26.3	5.0	808	7.6	6.1
1993/97	138	36.2	6.3	662	12.4	10.3

Note: unweighted statistics; "all" refers to the total number of sample observations in the balanced subpanel at times t and $t+s$.

Table 3: Transition Rates (Bivariate)

	ownership year t		ownership probability in $t+s$	
years $t/t+s$	stocks	mutual funds	stocks	mutual funds
1993/94	no	no	3.5	5.4
	yes	no	75.3	17.8
	no	yes	9.4	73.9
	yes	yes	81.3	72.9
1994/95	no	no	4.5	5.8
	yes	no	72.5	8.4
	no	yes	12.0	65.8
	yes	yes	74.6	70.9
1995/96	no	no	4.4	5.0
	yes	no	78.7	11.0
	no	yes	17.6	77.7
	yes	yes	79.1	83.6
1996/97	no	no	5.3	7.1
	yes	no	78.7	16.9
	no	yes	10.7	70.0
	yes	yes	82.5	79.6
1997/98	no	no	3.0	6.9
	yes	no	87.7	13.6
	no	yes	13.4	71.4
	yes	yes	80.8	76.9
1993/97	no	no	9.4	10.9
	yes	no	74.4	22.1
	no	yes	25.0	64.6
	yes	yes	73.8	61.9

Note: See Table 2

Table 4: Bivariate Random Effects Probit

	Specification 1				Specification 2			
	stocks		**mutual funds**		**stocks**		**mutual funds**	
Variable Name	***coeff.***	***s.e.***	***coeff.***	***s.e.***	***coeff.***	***s.e.***	***coeff.***	***s.e.***
constant	-4.8132	0.4400	-3.4065	0.3213	-4.7038	0.4299	-3.2691	0.3270
stocks (t-1)	1.1909	0.1077	-0.2540	0.1232	1.2983	0.1034	-0.1597	0.1250
mutual funds (t-1)	-0.1401	0.1073	1.1144	0.0946	-0.0534	0.1008	1.1305	0.0954
age	0.0210	0.0054	0.0091	0.0044	0.0204	0.0052	0.0087	0.0044
education								
intermediate	0.3623	0.1797	0.0491	0.1547	0.3052	0.1704	0.0249	0.1528
vocational	0.0372	0.1375	0.0225	0.1124	0.0322	0.1304	0.0225	0.1113
high	0.4054	0.1628	0.2161	0.1307	0.3517	0.1518	0.1897	0.1291
Wald (p-value)		*0.0028*		*0.2039*		*0.0076*		*0.3051*
log income	0.0522	0.0182	0.0219	0.0171	0.0500	0.0178	0.0205	0.0173
HH marg. tax rate	0.6158	0.3408	1.2520	0.2623	0.6180	0.3324	1.2343	0.2662
high-income panel	1.0030	0.1414	0.5981	0.1027	0.8546	0.1292	0.5358	0.0998
labor market status								
unemployed	0.1552	0.3431	-0.0121	0.2814	0.1420	0.3227	0.0302	0.2783
retired	0.3142	0.1519	0.3857	0.1274	0.2901	0.1457	0.3680	0.1265
disabled	-0.5144	0.3592	0.1424	0.2316	-0.4913	0.3386	0.1541	0.2313
selfemployed	1.5511	0.1695	0.1317	0.1538	1.4593	0.1660	0.1150	0.1559
other	0.4810	0.2371	-0.0329	0.1889	0.4570	0.2340	-0.0114	0.1873
Wald (p-value)		*0.0000*		*0.0487*		*0.0000*		*0.0794*
female	-0.4962	0.1536	-0.0736	0.1117	-0.4177	0.1485	-0.0482	0.1114
risk aversion								
low	---	---	---	---	0.3974	0.1398	-0.1125	0.1350
intermediate	---	---	---	---	0.2643	0.0807	0.0440	0.0687
don't know	---	---	---	---	-0.5045	0.3494	-0.8004	0.3000
not available	---	---	---	---	0.0940	0.1149	-0.2024	0.0940
Wald (p-value)		---		---		*0.0007*		*0.0048*
year								
1995	0.0614	0.0855	-0.0661	0.0739	0.0538	0.0844	-0.0728	0.0744
1996	0.2101	0.0966	0.0483	0.0904	0.1882	0.0943	0.0547	0.0910
1997	0.3542	0.1090	0.2106	0.0872	0.3154	0.1081	0.2142	0.0870
1998	0.6144	0.1378	0.3089	0.1039	0.5573	0.1418	0.3273	0.1050
Wald (p-value)		*0.0001*		*0.0021*		*0.0011*		*0.0014*
std.dev. RE	1.4446	0.1602	1.2027	0.1241	1.2506	0.1503	1.1294	0.1228
correl. RE	0.6590	0.0549			0.5994	0.0613		
correl. error	0.0260	0.0653			0.0791	0.0625		
# households	2861				2861			
# observations	9680				9680			
log-likelihood	-5706.42				-5664.56			

predicted probabilities at (st.(t-1), m.f.(t-1))	**stocks**	**mutual funds**	**stocks**	**mutual funds**
(0,0)	9.322	11.872	8.129	10.953
(0,1)	8.062	31.957	7.639	31.575
(1,0)	26.000	8.951	27.905	9.095
(1,1)	23.482	26.391	26.796	27.912

Note:

Reference groups for dummy variable groups are: low education; paid employee; high risk aversion. Estimates of the initial conditions equations are available upon request. The predicted probabilities are ownership rates (in %) obtained from the model for an "average" household: exogenous variables are set to their weighted sample means, random effects are set to zero, and the lagged ownership dummies are set to 0 or 1.

Table 5: Bivariate Random Effects Probit

	Specification 3				Specification 4			
	stocks		**mutual funds**		**stocks**		**mutual funds**	
Variable Name	***coeff.***	***s.e.***	***coeff.***	***s.e.***	***coeff.***	***s.e.***	***coeff.***	***s.e.***
constant	-12.5107	1.6092	-9.4148	1.0226	-11.4330	1.5161	-8.7041	1.0014
stocks (t-1)	1.1531	0.1575	-0.4358	0.1831	1.2681	0.1563	-0.3185	0.1774
mutual funds (t-1)	-0.4175	0.1630	0.9944	0.1415	-0.0967	0.1636	1.0613	0.1410
age	0.0008	0.0074	-0.0124	0.0062	0.0038	0.0071	-0.0119	0.0060
education								
intermediate	-0.0216	0.2592	0.0152	0.1984	-0.0641	0.2468	0.0066	0.1912
vocational	0.0307	0.2001	-0.0268	0.1508	0.0026	0.1884	-0.0413	0.1461
high	0.2577	0.2305	0.0789	0.1738	0.1783	0.2171	0.0208	0.1667
Wald (p-value)		*0.4469*		*0.8696*		*0.5908*		*0.9546*
log income	0.0824	0.0267	0.0123	0.0230	0.0838	0.0261	0.0151	0.0229
HH marg. tax rate	0.1087	0.4680	0.9149	0.3561	0.0884	0.4571	0.8680	0.3528
high-income panel	0.1288	0.1454	-0.0711	0.1195	0.1098	0.1376	-0.0743	0.1139
labor market status								
unemployed	0.5903	0.4752	-0.1068	0.4148	0.5987	0.4300	-0.0376	0.4034
retired	0.4507	0.2079	0.4781	0.1789	0.3791	0.2001	0.4209	0.1721
disabled	-0.1927	0.5550	0.2038	0.3207	-0.2193	0.5271	0.1765	0.3054
selfemployed	1.3236	0.2428	-0.4412	0.2180	1.2807	0.2400	-0.4284	0.2109
other	0.5131	0.3815	0.2807	0.3226	0.4569	0.3835	0.2819	0.3188
Wald (p-value)		*0.0000*		*0.0219*		*0.0000*		*0.0357*
female	-0.1328	0.1794	0.2111	0.1520	-0.1211	0.1738	0.1794	0.1477
log fin. wealth (t-1)	-0.1089	0.0504	-0.0925	0.0554	-0.1227	0.0526	-0.1023	0.0537
log fin. wealth (avg.)	0.9535	0.1235	0.8048	0.0992	0.8445	0.1166	0.7524	0.0972
risk aversion								
low	---	---	---	---	0.5097	0.1933	-0.1713	0.1912
intermediate	---	---	---	---	0.3635	0.1184	0.0779	0.0875
don't know	---	---	---	---	-0.3167	0.5151	-0.5769	0.4641
not available	---	---	---	---	0.3147	0.1603	-0.0479	0.1336
Wald (p-value)		---		---		*0.0091*		*0.3394*
year								
1996	0.1406	0.1065	0.1288	0.1006	0.1122	0.1060	0.1336	0.0997
1997	0.2737	0.1294	0.3247	0.1012	0.2226	0.1314	0.3164	0.1012
1998	0.5503	0.1705	0.4428	0.1247	0.4544	0.1749	0.4331	0.1272
Wald (p-value)		*0.0123*		*0.0011*		*0.0721*		*0.0022*
std.dev. RE	1.4101	0.2223	1.1605	0.1787	1.2051	0.2112	1.0361	0.1760
correl. RE	0.6074	0.0995			0.4247	0.1214		
correl. error	-0.1169	0.0910			0.0208	0.0910		
# households	1871				1871			
# observations	5953				5953			
log-likelihood	-3401.50				-3383.59			
predicted probabilities at (st.(t-1), m.f.(t-1))	**stocks**		**mutual funds**		**stocks**		**mutual funds**	
(0,0)	5.638		7.440		4.713		6.739	
(0,1)	3.382		21.339		4.137		22.407	
(1,0)	17.909		4.197		19.392		4.302	
(1,1)	12.295		14.026		17.742		16.362	

Note: see Table 4

Table 6: Linear Probability Models

	Stocks		Mutual Funds	
Variable Name	Estimate	Std.Err.	Estimate	Std.Err.
constant	–0.1403	0.0540	–0.1368	0.0566
stocks_{t-1}	0.1734	0.0541	0.0102	0.0503
$\text{mutual funds}_{t-1}$	–0.0640	0.0392	0.2044	0.0494
age	0.0025	0.0009	0.0035	0.0009
education				
intermediate	0.0603	0.0270	0.0063	0.0249
vocational	0.0148	0.0178	–0.0028	0.0171
high	0.0747	0.0259	0.0442	0.0243
log income	0.0073	0.0034	0.0015	0.0032
HH marg. tax rate	–0.0436	0.0444	0.1284	0.0621
high–income panel	0.1439	0.0211	0.0849	0.0212
labor market status				
unemployed	0.0070	0.0290	0.0230	0.0283
retired	0.0476	0.0293	–0.0263	0.0321
disabled	0.0257	0.0386	0.0048	0.0395
self–employed	0.2239	0.0580	0.0105	0.0369
other	0.0461	0.0232	–0.0197	0.0260
female	–0.0395	0.0160	–0.0164	0.0170
year				
1996	0.0176	0.0074	0.0033	0.0077
1997	0.0347	0.0093	0.0365	0.0109
1998	0.0638	0.0124	0.0539	0.0147
Number of households	1870			
Number of observations	5950			
Sargan, 14df ($p-$value)	26.7231	*0.021*	19.1600	*0.159*
AR(2) test ($p-$value)	0.094	*0.925*	1.336	*0.181*

Table 7: Oaxaca Decompositions of Changes in Ownership Rates (in percentage points)

years	1993/94	1994/95	1995/96	1996/97	1997/98	1993/97
Change in ownership rate			Stocks			
Total change	n.a.	1.40	1.87	1.85	3.48	n.a.
exlained by change in …						
stocks$_{t-1}$ & mutual fds.$_{t-1}$	n.a.	0.09	0.29	0.17	0.11	n.a.
education	0.00	0.04	0.00	–0.00	0.01	0.02
age	0.19	0.21	0.23	0.24	0.26	1.10
log income	0.16	–0.02	–0.02	–0.22	0.16	–0.02
tax rate	0.08	–0.03	–0.02	–0.12	0.06	–0.20
tax & income	0.23	–0.07	–0.05	–0.36	0.20	–0.25
labor market	0.00	0.50	–0.02	0.02	–0.02	0.69
all x-s	0.42	0.70	0.18	–0.09	0.44	1.62
time dummies	0.00	0.60	1.57	1.61	3.03	n.a.
Change in ownership rate			Mutual Funds			
Total change	n.a.	–0.71	1.71	2.29	1.80	n.a.
exlained by change in …						
stocks$_{t-1}$ & mutual fds.$_{t-1}$	n.a.	0.27	–0.25	0.12	0.16	n.a.
education	0.00	0.04	–0.00	0.00	0.01	0.04
age	0.12	0.12	0.13	0.14	0.14	0.60
log income	0.07	0.00	0.00	–0.10	0.07	0.03
tax rate	0.13	–0.08	–0.10	–0.34	0.15	–0.56
tax & income	0.18	–0.09	–0.11	–0.47	0.20	–0.59
labor market	0.00	0.18	0.12	0.14	0.06	0.65
all x-s	0.29	0.25	0.13	–0.18	0.43	0.74
time dummies	0.00	–0.88	1.55	2.35	1.49	n.a.

Note: Figures in this Table derive from Specification 1 in Table 4. They are based on weighted means of (univariate normal) ownership probabilities as predicted from equations (1) and (2), with random effects set to zero; presented are the differences in such means between the baseline case and the case where some regressors are lagged: "total change": all right hand side variables are lagged, including ownership dummies and time dummies; "all x-s": all regressors except time dummies and lagged ownership dummies are lagged.

Figure 1: Ownership by Cohort: Stocks

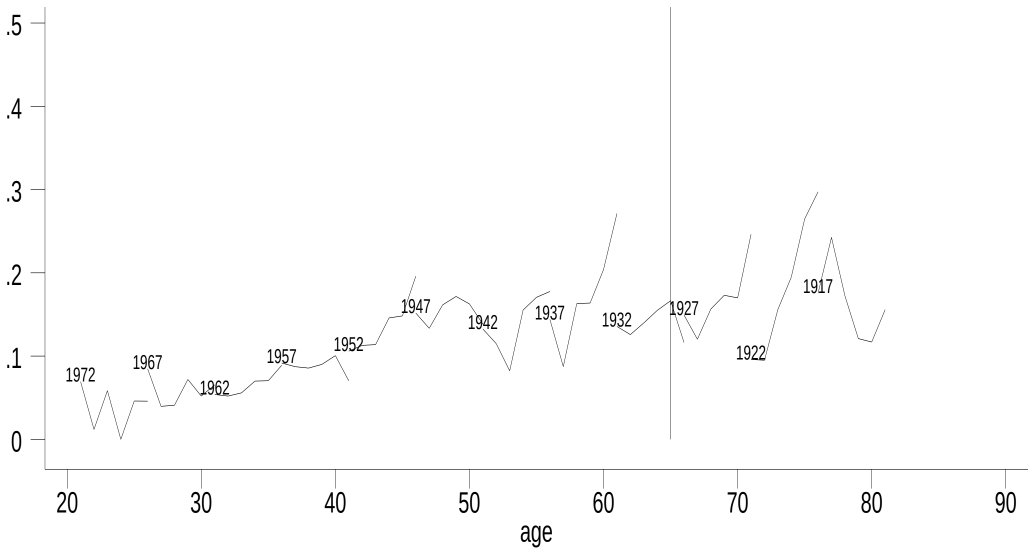

Figure 2: Ownership by Cohort: Mutual Funds