Household Portfolio Allocation in the Netherlands: Saving Accounts versus Stocks and Bonds

by

Stefan Hochgürtel,
Rob Alessie,
Arthur van Soest,*†
Tilburg University

February, 1995

Abstract

This paper analyzes the portfolio structure of households in the Netherlands. It considers the allocation of financial wealth to two major asset categories, namely saving accounts on the one hand and stocks and bonds on the other hand. The latter category is considered to be more risky than the former. We analyze the impact of the overall level of wealth, the marginal tax rate, and other variables on the allocation between assets, using cross–section data drawn in 1988 that provide detailed information on the structure of household wealth, not only on ownership but also on the amounts of wealth held in the respective asset categories. The econometric specification is a trivariate tobit type model. One equation explains the total level of wealth, a second one explains individual threshold values below which no wealth is held. The third equation explains the share of wealth invested in stocks and bonds. The model is estimated using Full Information Maximum Likelihood. Limited information provided by the data (non–reporting) is explicitly taken into account. Results show that wealth and the marginal tax rate are major determinants of the allocation between safe and risky assets.

Keywords: household saving, portfolio choice, limited dependent variables

JEL classification: *C34, D12, G11*

*correspondence address of the authors:
CentER for Economic Research, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands; e–mail: S.Hochguertel@kub.nl, Rob.Alessie@kub.nl, A.H.O.vanSoest@kub.nl

†We are grateful to Research International Nederland (RIN) for providing the data, and to an anonymous referee for helpful suggestions. All remaining errors are ours. The first author is supported by the VSB savings project, research of the third author is made possible by a fellowship of the Netherlands Royal Academy of Arts and Sciences (KNAW). A previous version of this paper has appeared under the same title as VSB–CentER Savings Project Progress Report, No. 23, Tilburg University.

I. Introduction

This paper deals with the structure of household portfolios in financial wealth in the Netherlands. The question of how people allocate their portfolios has attracted considerable interest in the past, from a theoretical, empirical and economic policy point of view. One of the central questions is whether the portfolio structure has an impact on the overall savings rate and whether the government can influence national saving — which is a major determinant of growth and the capability of the economy to invest — by means of differential tax treatment (see the discussion on tax–deferred saving incentives in e.g. Venti and Wise (1986) or, more recently, Engen et al. (1994)). Moreover, the portfolio allocation determines the risk the economy is willing to bear, which can have substantial impacts on welfare, and can itself be influenced by policy instruments. For instance, policies related to the distribution of wealth can have an impact on portfolio composition.

The motivation for most of the empirical literature on household portfolio choice, stems from a discrepancy between theoretical predictions and observed outcomes. For the analysis of portfolio choice, the standard model in financial economics is the capital asset pricing model (CAPM). It implies that investors only care about mean and variance of their assets. In equilibrium they will hold only the minimum–variance portfolio (if a riskless asset is available this will be the only asset in that portfolio) and the market portfolio (which is a value–weighted mutual fund of all securities available in the market) independently of their tastes and endowments. The only difference between individuals is the proportions invested in the two portfolios. According to their individual risk tolerance, they either hold a combination of the riskless asset and the market portfolio or they leverage their portfolio by holding a short position in the riskless asset by borrowing (see the recent, non–technical overview in Sharpe (1991)).

In the empirical literature however, evidence is found that individuals hold portfolios of very diverse nature, usually containing only a small number of different assets. To explain this considerable lack of diversification per individual, fixed transaction costs in the form of holding or monitoring costs have been considered (cf. Goldsmith (1976) and Mayshar (1981)). The idea behind this is that the net return on assets depends on the amount held in the asset. Furthermore, institutional restrictions such as minimum purchase requirements that act as rationing schemes or constraints on short sales induce consumers not to hold particular assets (cf. Deaton (1981) for the impact of rationing on demand). Yet another approach deals with explaining why the portfolio structure is biased towards or against certain assets without necessarily leading to zero holdings. This includes the analysis of taxes and differential tax treatments which alter the relative price of assets (Feldstein (1976)) or capital market imperfections such as borrowing or liquidity constraints (Paxson (1990)) or incomplete information (King and Leape (1987)).

The empirical literature on household portfolio choice based on survey data commences with a study by Uhler and Cragg (1971) who use a multinomial logit model to explain the level of diversification. They analyze the effect of wealth and disposable income, as well as age and family size, on the choice of the level of diversification. The idea is that fixed transaction costs become less important as the total volume of resources increases. Alessie et al. (1993a) center on the endogeneity of marginal tax rates in the choice of the portfolio, which arises from different tax rates for different assets. Thus, the level of marginal tax rates induces further diversification which in turn affects taxable capital income. Taxable income is thus treated as an endogenous variable in the problem of the number of assets chosen. King and Leape (1987) and Ioannides (1992) focus on the number of asset categories in a household asset portfolio as well.

Feldstein (1976) estimates shares equations for different assets, stressing the importance of taxes in the problem of portfolio choice, but neglects all zero wealth observations. Similar in method, but different in emphasis is the paper by Hubbard (1985), who analyzes the influence of holding social security wealth and private pension wealth on asset demand, conditional on these assets being held at all. King and Leape (1984) stress that the portfolio choice problem involves both a decision which assets to hold and how to allocated resources between assets; these two decisions are inseparable. They investigate the diversification problem by estimating a switching regression model with endogenous switching. The choice problem consists of a discrete choice of which assets to choose, and an asset demand function conditional on ownership. Agell and Edin (1990) combine features of the King and Leape and Feldstein papers by specifying asset demand functions, conditional on ownership, as a function of household characteristics, where a dummy indicates the chosen portfolio; since these dummies are considered endogenous, they are replaced with predicted probabilities from separate, reduced form probits per asset ownership combination. In each conditional asset demand equation — they are all estimated separately — features a Heckman correction term from univariate probits of asset ownership to correct for sample selection bias. The marginal tax rate as right–hand–side variable is instrumented to allow for endogeneity. Guiso et al. (1994) set up a single equation two–limit tobit model for the share of various risky assets. Their major concern is the impact of borrowing constraints and uninsurable income risk on the portfolio structure. Additional empirical investigations have been conducted by Shorrocks (1982), King and Dicks–Mireaux (1982) and Ioannides (1992). Amemiya et al. (1993) compare three tobit estimators for demand equations of two assets. The first arises from a maximizing quadratic utility in assets under non–negativity constraints. The second explicitly models the impact of transaction costs on the level of diversification and explains conditional on that, individual asset demands. The third estimator models the investment decision as a sequential choice where demands for two assets are generated from the residual wealth after the demand for a first asset has been subtracted.

From the economic modelling point of view, most related to the idea of the present paper is probably Venti and Wise (1986). These authors estimate demand equations derived from an objective function with branch structure. They distinguish between income allocation between consumption and savings on one level, and between assets on the other level.

The present paper wants to shed some light on the portfolio structure of households in the Netherlands. We consider the allocation of wealth between two major asset categories, namely saving accounts on the one hand, and stocks and bonds on the other hand. The latter category is considered to be more risky than the former. We use a data set that provides detailed information on the structure of household wealth, not only on ownership, but also on the amounts of wealth held in the respective asset categories. We consider a three equation model. The first equation is a censored regression equation in which total wealth invested in savings or stocks and bonds (to be referred to as wealth in the sequel) is explained from family income and other background variables. By means of a second equation we take into account the possibility of individual differences in a lower bound for demand for either one of the two assets. The third equation is a two–limit tobit equation in which the share of total wealth invested in stocks and bonds is explained from the level of wealth and from family characteristics. We allow for endogeneity of wealth in the share equation and for selection bias due to zero wealth or unobserved asset amounts. The latter is possible since the routing of the questions allows us to infer that the amount is positive whenever no amount is given but the ownership question has been answered in the affirmative.

The paper is organized as follows: The next section gives a short economic motivation for the econometric model, whose structure is outlined in the succeeding section. An overview of the data that have been analyzed is given in section IV. Results are discussed and evaluated in section V. The concluding section shortly discusses possible extensions of the model into several directions.

II. Economic Motivation

To motivate the empirical analysis in this paper, we formulate a simple one–period model for each household. At the beginning of the period, the household can choose how to allocate the available amount Y, consisting of, say, assets available from the previous period and current income. It can allocate Y in various ways: consumption, durables, housing, illiquid assets, and financial assets.[1] We are interested in two types of financial assets, one riskfree and one risky. Let q_A denote the vector of amounts allocated to these two assets, and let q_B denote the vector of amounts allocated in other ways. Because of consumption and because of the (distribution of the) value of the assets at the end of the time period, each allocation has some (expected) utility level $U(q_A, q_B)$. We assume that the household maximizes utility subject to the budget allocation constraint:

$$\max_{q_A,q_B} \quad U(q_A, q_B)$$

[1] Loans could easily be incorporated as well, with negative utility.

$$\text{s.t.} \qquad (q_A, q_B)'\iota = Y$$

Here, ι is a vector of ones of appropriate dimension. For the general utility function given above, the optimal quantities of interest q_A will be a complicated function of all variables determining utility of q_A and q_B, i.e., for example, prices of consumption and housing. We want to focus on the allocation of financial wealth between different types of assets. The natural way to achieve this is to impose restrictions on the preference structure, i.e. to assume direct weak separability:

$$U(q_A, q_B) = U(U_A(q_A), q_B)$$

Here $U_A(q_A)$ could, for example, be given by

$$U_A(q_A) = \mathrm{E}[u(R'q_A)]$$

where R is the two–dimensional random vector of pay–offs of the two financial assets, and where u is some utility function, concave in case of risk aversion. The assumption of direct weak separability allows for two stage budgeting (see for instance Deaton and Muellbauer (1980), Phlips (1983) or Blackorby et al. (1978)): the utility maximization problem can be written in two steps. The first step (i.e. the second stage of the budgeting process) is

$$\begin{array}{rl} \max_{q_A} & U_A(q_A) \\ \text{s.t.} & \iota' q_A = W \end{array}$$

for a given total financial wealth allotment W. From its solution, $q_A^\star(W)$, the indirect utility function $V_A(W) = U_A(q_A^\star(W))$ obtains. The second step concerns the allocation of W between the two assets and may be represented by the problem

$$\begin{array}{rl} \max_{q_B, W} & U(V_A(W), q_B) \\ \text{s.t.} & \iota' q_B + W = Y \end{array}$$

from which the optimal values $q_B^\star, W^\star, q_A^\star = q_A(W^\star)$ follow. The goal of this kind of modelling is to focus on structural financial portfolio decisions that can be separated from other issues. It is assumed that decisions on q_A, conditional on W, can be treated independently of the decisions on q_B. A motivation for this approach is that financial assets are different from other wealth categories which comprise mainly housing and social security wealth. Weak separability in the preference structure is a necessary and sufficient condition for the second stage of the budgeting process: maximizing a weakly separable utility implies that each sub–utility must be maximized subject to the respective sub–budget; conversely, the existence of subgroup demand functions implies weak separability (cf. Deaton and Muellbauer (1980)).

In our empirical model, we explain W and, conditional upon W, we explain q_A. We first discuss the latter. According to the second step, q_A is determined by W, by the after tax pay–offs of the riskfree and the risky asset, and by taste shifters that determine the attitude of the household towards risk. Note that net returns on the assets vary across households, since the tax system implies that the marginal tax rate on the returns depends on family composition and incomes of family members. We directly allow the utility function U_A to vary with the household's marginal tax rate. To allow for preference heterogeneity, we include several other family characteristics. These are, together with a constant term and the marginal tax rate, included in a vector x_s. We allow for an error term ϵ_s reflecting unobserved heterogeneity. We choose the following specification:

$$s = x_s\beta_s + \gamma \log W + \epsilon_s \tag{1}$$

Here s is the share of W invested in stocks and bonds. The share invested in saving accounts is given by $1 - s$ and has the same form. Equation (1) corresponds to the Engel curves of the Almost Ideal Demand System (Deaton and Muellbauer (1980)) or the translog demand system (Christensen et al. (1975)).[2] The case where $W = 0$ can be handled as well, since the equation is only used for individuals with positive wealth.

The Engel curve specification given above does not yet allow for non–negativity constraints. Such constraints — due to constraints on short sales, for instance — that lead to different regimes in the demand equations can be included as additional optimization restrictions. Since we are dealing with only two assets, incorporating non–negativity constraints is straightforward and leads to corner solutions on the budget line.

The first stage concerns the choice of W. This basically depends on everything else: pay–offs of financial assets, pay–offs and direct utility of other assets such as housing or durables, initial wealth, current and future income, etc. We do not aim at developing a structural model for W. Instead, we formulate two reduced form equations for W in the next section. W is explained from a number of background variables and observed current family income. Since unobserved heterogeneity will in general affect both the optimal share and total wealth, we allow for correlated error terms across equations, i.e. W can be endogenous in the shares equation.

III. Econometric Model

Equation (1) determines the allocation of financial wealth W between saving accounts on the one hand and stocks and bonds on the other hand. We thus estimate a single demand equation for the share of stocks and bonds s. Moreover we add two equations to explain W. The wealth equation is a reduced form equation, in which wealth is explained from income and socio–demographic control variables. To explain the fact that those people who hold financial wealth

[2] In preliminary estimates we also included $\log^2 W$ in (1). This leads to the quadratic system of, for example, Banks et al. (1994). The quadratic term however, was insignificant.

$(W > 0)$, usually hold a substantial amount, we add a threshold equation for a minimum value of W, following Nelson (1977). The latter is included in order to increase the empirical content of the estimated model. The threshold equation contains the same right–hand–side variables as the wealth equation.

As already stated in the previous section, we shall work with a logarithmic transformation of wealth: $w^\star = \log W$. We write $w^\star$ and not w to indicate that, in case of incomplete reporting, this variable can be unobserved. For the same reason, the actual share of W invested in stocks and bonds is denoted by $s^\star$. The threshold value $T^{\star\star}$ is always unobservable. To control for endogeneity of the level of wealth, and to take account of the fact that one or two of the amounts invested are sometimes unobserved, we estimate the three equations jointly.

We first model latent variables $w^{\star\star}$, $T^{\star\star}$ and $s^{\star\star}$:

$$\begin{array}{rcllllll} w^{\star\star} & = & x_w\beta_w & + & & & \epsilon_w \\ T^{\star\star} & = & x_T\beta_T & + & & & \epsilon_T \\ s^{\star\star} & = & x_s\beta_s & + & w^\star\gamma & + & \epsilon_s \end{array} \qquad (2)$$

Here $w^{\star\star}$ can be seen as a measure of desired wealth, which may be negative. $s^{\star\star}$ is the desired fraction of stocks and bonds in wealth. It may be less than zero or larger than one. The error terms $(\epsilon_w, \epsilon_T, \epsilon_s)$ are assumed to be trivariate normally distributed with mean zero and independent of x_w, x_T and x_s. Their covariance matrix is

$$\Sigma = \begin{pmatrix} \sigma_w^2 & \rho_{wT}\sigma_w\sigma_T & \rho_{ws}\sigma_w\sigma_s \\ \cdot & \sigma_T^2 & \rho_{Ts}\sigma_T\sigma_s \\ \cdot & \cdot & \sigma_s^2 \end{pmatrix}.$$

Actual wealth may differ from desired wealth because individuals may not find it rational to invest in the considered wealth category unless the amount is substantive:

$$w^\star = w^{\star\star}\mathbf{1}_{[w^{\star\star} > T^{\star\star}]},$$

where $\mathbf{1}_{[\cdot]}$ is an indicator which is one if the argument is true, and zero otherwise. We assume that the probability of the threshold falling below zero is small so that the possibility that $W^{\star\star}$ can be ignored (see the estimation results below).

The threshold equation, as well as the wealth equation, is modelled in a reduced form way. A structural approach along the lines sketched in section II would be very complicated and requires more information than we have in our data. Thus, the first two equations reflect both demand and supply side factors which determine the equilibrium in the market for financial assets. An interpretation of this threshold equation is the measurement of (unobserved) fixed transaction costs which are imposed in a lump–sum way and influence the decision whether or not to hold wealth in the assets under consideration. We would expect these costs, such as monitoring costs, to be individual specific since they might be (directly or indirectly) a function of observables like income or family characteristics.[3] [4] Secondly, the thresholds may reflect institutional restrictions such as minimum purchase requirements for stocks and bonds or effective non–negativity constraints on saving accounts.[5]

The shares equation is only defined for cases in which $w^{\star\star} > T^{\star\star}$. The actual share in stocks and bonds, $s^\star$, may differ from the desired share $s^{\star\star}$ because of the non–negativity constraints on both assets:

$$s^\star = \begin{cases} 0 & \text{if} \quad s^{\star\star} \leq 0 \\ s^{\star\star} & \text{if} \quad 0 < s^{\star\star} < 1 \\ 1 & \text{if} \quad s^{\star\star} \geq 1 \end{cases}$$

We abstract from underreporting or other forms of misreporting. Thus, if the amounts invested in both assets are observed, total wealth W and $w = w^\star = \log W$ are observed. It happens rather often however, that respondents — although they indicate to own the respective assets — refuse to give the amount (in 354 out of 2849 cases, cf. section IV). In this case, the only information we have is that $w^\star > T^{\star\star}$ whereas the magnitude w is missing.

If both amounts are reported, we also observe the share $s = s^\star$. If one of the amounts is positive but not reported, we observe $s^\star$ if and only if the second amount is zero. In that case $s^\star$ is either zero or one. Due to non–reporting in at least one of the asset categories, $s^\star$ is unobserved for 344 out of 2658 observations.

The model used here is basically of a recursive simultaneous tobit type (cf. e.g. Maddala (1983)). Wealth has a lower unobserved threshold value for each individual and the fraction of stocks and bonds is bounded in the interval $[0, 1]$. The model does not allow for negative values of wealth or values of the share outside the interval $[0, 1]$. Moreover, it allows for a positive probability that wealth is exactly zero and for positive probabilities that the share is exactly zero or exactly one. Conditional on being positive, wealth has a continuous distribution, and conditional on being in the open interval $(0, 1)$, the share follows a continuous distribution.

It should be stressed that the model is specifically tailored to the information given in the data: the double starred variables are latent magnitudes for partly unobserv*ables* whereas the

[3]An indication of switching or portfolio adjustment costs is provided in the data by a variable that measures the frequency of changes in the portfolio of stocks and bonds (excluding options): 60 % of the households holding these assets make adjustments about once a year or even more seldom.

[4]Agell and Edin (1990) argue against a simple tobit specification in the shares equation if asset demand functions conditional on ownership differ from the function for the discrete portfolio choice: fixed transaction costs only affect the discrete portfolio choice problem but not the conditional asset demands, since they are sunk (Mayshar (1981)). We remedy the empirical shortcoming of this kind of modelling by allowing individual thresholds, reflecting at least partly the presence of transaction costs, to influence the decision to hold wealth in financial assets.

[5]Nelson (1977, fn. 6) notes that the model would be different if all the cited possibilities for the existence of thresholds work simultaneously but in differing directions such that a model similar to a switching regression model would be appropriate.

single starred symbols are in principle observable but in fact partly unobserv*ed*. Threshold values are unobservable throughout. Ignoring the information for unobserved wealth cases — which would be a "natural" first approach — would lead to biased and inconsistent parameter estimates, as would ignoring households that do not hold the assets at all. The reason is that wealth can only be unobserved if it is positive, so that observability of wealth and level of wealth are automatically correlated. The present framework avoids these traps since it makes use of all observations in the sample.

To complete the specification of the model, we have to add an assumption on what determines whether a positive amount invested in one of the two asset categories will be observed. We assume observability of asset amounts, conditional on the fact that the assets are held, to be exogenous, i.e. the response is independent of the wealth held.[6]

Note that, although the model does take the problem of non–reporting appropriately into account, we have to rely on the assumption that we do not face problems with measurement errors (due to under–reporting for instance) with possibly non–zero expectation.

The estimation proceeds then as a full information maximum likelihood estimation which allows to take the correlation structure between the three equations in (2) fully into account. With wealth composed of two assets, nine regimes can be derived: each of the two assets can be zero, positive and observed, or positive and unobserved. Each of these regimes requires its own expression for the likelihood contribution. These likelihood contributions are noted in the appendix. Given the distributional assumption, the estimator is consistent and asymptotically efficient. Identification of the model is achieved by setting both covariances between ϵ_w and ϵ_T and between ϵ_T and ϵ_s to zero and by excluding at least one variable in $x_w \equiv x_T$ from x_s.

IV. Data

In the analysis below we use data which provide a rich source of information on the structure of individual wealth, which is much more detailed than other comparable data from the Netherlands.

The Dutch Collective Bank Study (Collectieve Banken Onderzoek, CBO), sampled in 1988, consists of 3704 households with 10113 individuals in the Netherlands. The topic of the survey is targeted at the financial structure of household and individual wealth and the institutional

[6]To be more precise, let us assume that observability of the two assets $(j = 1, 2)$ is determined by the following equations:

$$O_j^\star = X_j \beta_j + \epsilon_j \qquad (j = 1, 2)$$

Here, $O_j^\star$ indicates whether the individual is prepared to answer the questions on the amount in asset j ($O_j^\star > 0$) or not ($O_j^\star < 0$). This equation is irrelevant for people who do not hold the asset. We assume independence of the error terms (ϵ_1, ϵ_2) from $(\epsilon_w, \epsilon_T, \epsilon_s)$. This allows to drop the observability equations in estimation, since in the maximum likelihood procedure that we employ, the log–likelihood can be written as a sum of the log–likelihood of these two additional equations, and the log–likelihood of the equations of interest.

relationships of consumers to banks and financial intermediaries.

In each household the head of household (and, if present, his or her partner and children below 18), and other household members above 18 were interviewed. The questionnaire comprises detailed information on general financial behavior, saving accounts, checking accounts and credit cards, stocks and bonds, loans, mortgages, and insurances, as well as some background information on the family structure, employment status, net income and socio–demographic characteristics.

The section on savings deals with the possible types of saving accounts, amounts as well as increments and decreases in savings, and motives for savings. In addition, there are sections on fixed time deposits and saving certificates and certificates of deposit. A further section contains questions on stocks and bonds, which include the wealth held in these assets and the bank relationships.

One of the characteristic features of the data set is the detailed information on single asset units: the amount and the institution (bank) were requested for each single account. This is a unique feature compared with other data sets that only give crude information on asset amounts, if at all. This leads us to believe that the information contained in the CBO data is much more reliable and qualitatively more valuable than in other data sets. Moreover, from aggregate data it is known, that this data set resembles, at least as far as saving accounts are concerned, the national figures better than the SEP (Sociaal Economisch Panel). In particular, ownership rates of saving accounts are much better represented in the CBO (Alessie et al. (1993*b*)). In both CBO and SEP however, there seems to be considerable underreporting on the amounts.

The survey was carried out by Research International Nederland (RIN). Within each asset category ownership questions had to be answered and, conditional on this being positive, the amount was requested. In the questions regarding the ownership we have virtually no missing values as opposed to the variables regarding the amounts.

Since this paper deals with the structure of wealth on household level, the individual responses had to be aggregated to household level. This resulted in 3704 household observations. We thus assess the financial situation of the entire household, which is assumed to form one single decision unit. Amounts for specific assets have been summed over all assets in the asset category and over all household members. It has been assumed that the household characteristics for non–wealth related variables can be replaced by the respective characteristics of the head of the household. In terms of the definition of the head of the household we largely follow the conventions of the SEP.[7]

[7] The SEP defines the head of the household as the husband of a married couple or the parent for single parent households or — apart from few exceptions — the person that is considered to be head by other household members. The self–reported indicator for the position within the household was not considered as identifier for the head of the household. By the assignment used here it is ensured that each and every household has one and only one head.

In the estimated model 501 observations dropped out due to missing values in the explanatory variable on net income, an additional 32 observations were discarded as severe outliers in this respect, leaving 3171 observations. However, comparing statistics of these data with the respective statistics from the full sample shows, that the distribution of the observations used resembles the full sample distribution for all variables quite neatly. This suggests that selection bias due to unobserved income is not a serious problem. The distribution of financial wealth is dominated by saving accounts (cf. table 1). The savings variable comprises the amounts held in saving account balances, time deposit accounts and saving certificates and certificates of deposit. 84% of all households report to have positive amounts of savings, whereas only 200 households (6.3%) hold stocks and bonds. The latter variable is the sum of shares in domestic companies, shares in foreign companies, shares in investment fonds, options, bonds and mortgage bonds. Only 191 (6%) of the observed households hold both asset categories. 39 households holding stocks and bonds refused to give the value of their stocks and bonds, 315 households with positive savings did not report the amount. These include 10 households that reported neither of the two amounts, although having answered both ownership questions in the affirmative.

Summary statistics of explanatory and dependent variables are presented in table 2. The distribution of stocks and bonds is highly skewed with a skewness[8] parameter of 16.3. The mean of all reported stocks and bonds is about 3800 Dfl., the mean of the positive observations is about 73,000 Dfl. For savings the picture is less dramatic: the mean of all observed saving accounts balances is around 13,000 Dfl., while the conditional mean of positive observations is about 16,000 Dfl. The skewness amounts to 10.6. 18.1% of the wealth observations are strictly positive, the mean and median for financial wealth, conditional on being positive, are about 20,000 and 4,700 Dfl., respectively.

Due to non–reporting, direct information on total wealth invested in the two assets is available in only 2827 cases. Looking at ln(wealth + 1), a variable which is closely related to the one used in the empirical analysis (ln(wealth)), except that it includes the cases where wealth is zero, has a mean of 6.8, whereas its median is 8.0. The income variable (ln(income + 1)) used in the analysis has mean and median both equal to 7.74. The skewness parameters of ln(wealth+1) and ln(income+1) are negative: -0.92 and -2.61 respectively, whereas wealth and income themselves are skewed to the right (skewness parameters 11.5 and 19.1, respectively).

Turning to the background variables, from table 2 it can be seen that 59.8% of the heads of the households were full–time employed, and only 4.9% part–time. This is partly due to the fact that only 20.2% of the households were headed by women. Four dummies on the occupational status of the head of the household are used (farmer, self–employed or manager, white–collar, blue–collar), covering 79.7% of all households. The remaining 20.3% are mainly nonworkers. Looking at the geographical distribution, 19.6% of all households were living in the southern

[8] as defined by skewness$(x) \equiv \mathrm{E}(x - \mathrm{E}(x))^3/\sigma^3$, where σ^2 is the variance of x.

provinces (Noord–Brabant and Limburg), whereas 34.1% lived in five northern regions. The reference category includes the provinces in the highly populated west of the country. The other variables in table 2 are self–explanatory.

To get some rough ideas about the extent to which Dutch households diversify their portfolios, tables 3 and 4 provide information on the total level of diversification in the data. The assets in the data set can be aggregated to 4 broader asset categories as: saving accounts and related assets, stocks and bonds, housing wealth and life insurance contracts. The figures show that 30% of all households in the sample hold only one asset, 34% hold two, 25% three, and only 3% of the households hold a "fully" diversified portfolio. The mean number of asset categories is around 1.83 whereas the median is 2, the standard deviation amounts to 0.99. From table 3 it can be seen that no households exclusively hold stocks and bonds. Out of the 200 households who do hold stocks and bonds, 106 also have a life insurance. Moreover, of those holding stocks and bonds, 80% own a house. The picture is a bit different for families possessing saving accounts: 47% of these households have a life insurance policy, and 53% own a house. The data reveal a substantial correlation between house and life insurance ownership. 56% of house owners have a life insurance policy, 61% of life insurance bearers own their house. Two things are worth mentioning: First, there is considerable lack of diversification within the groups, i.e. most households hold only very few assets within each category. Second, even the full range of assets covered by the questionnaire is not a complete compilation of all assets in the market.

Finally, the survey provides information on other wealth–related magnitudes such as mortgages and loans, durable goods, insurance contracts and checking accounts. These categories have not been considered here. Checking accounts are usually viewed as a temporary storage for incoming resources and source for daily expenditures, and thus are more a reflection of the flow of income and saving rather than a stock of a particular wealth category. Moreover, virtually every household (99%) in the sample does hold checking accounts. Insurance contracts do only indirectly contribute to wealth, durables are to a large extent consumption goods, and mortgages and loans do contribute to the net but not to the gross wealth position of the household.

V. Results

The estimation results are presented in table 5. The parameters give the change in the expected value of the dependent variables, i.e. $w^{\star\star}$, $T^{\star\star}$ or $s^{\star\star}$, in response to a change in the exogenous variables. The left hand panel of the table contains the estimates for the wealth equation, i.e. the total amount invested in the two assets. From this it can be seen that (net) income has a strong and significant effect on the level of wealth. The income elasticity of wealth is 0.47. Evaluating the estimated equation at sample means reveals log income to be the major determinant of wealth.

The marginal tax rate on income is found to be highly significant and shows a considerable tax responsiveness of households to hold savings and stocks and bonds.[9] [10]

Age has a strong positive and significant effect on wealth. A squared age term has been included in order to capture non–linearities, although the first order term is insignificant. The estimated age function is increasing throughout for the sample observations. It should be noted that pure age effects cannot be disentangled from cohort effects in this cross–section study (for attempts in this direction: see King and Dicks–Mireaux (1982)). On average, full–time workers hold higher wealth than the reference group (which comprises employees who work only few hours, disabled persons and unemployed, retirees, students and housewives), which in itself is not surprising.

The level of education is controlled for by inclusion of a dummy variable for high education, which is negative but insignificant, and an interaction term for high education and age, which is positive and significant at the 5% level. Together these estimates suggest that wealth increases with educational level for all but the youngest age groups.

The parameter estimate for blue– and white–collar workers (although insignificant for the latter group) indicate a lower engagement in the considered asset groups, compared to the reference group (people without paid job and others). An explanation could be that these persons hold more wealth in other assets which are not included here.

In order to control for possibly different saving behavior of farmers, a dummy variable for this profession has been included and the estimate turns out to be significant, whereas other selfemployed persons or employees in a managerial position do not hold significant amounts in the assets under consideration. It should be noted however, that the questionnaire usually does not distinguish between private and business assets of the self–employed and farmers. In this respect, one would expect a positive effect of being self–employed if shares in own businesses are included in the amount given. Moreover, the decision of being a farmer and the decision of how much wealth to hold may be interrelated.

Family composition appears to be an important determinant of wealth holding: married or unmarried couples and parents with children hold significantly more wealth than single individuals. Savings for future children or for children that have already left the household, could explain why couples without and couples with children behave similarly. Finally, there is a geographical effect: households from northern regions hold more wealth on average than households from the west. An explanation might be that housing prices in the northern areas are lower than in other regions. If investment in housing is inelastic, this might increase the

[9] on 5% significance level; this is the level chosen throughout

[10] According to Agell and Edin (1990), the marginal tax rate might be endogenous since it applies to income out of wealth as well. Since we have constructed the marginal tax rate on basis of measured net monthly income, which probably excludes asset income, this problem might be less severe here. Due to lack of appropriate instruments, we could not test for endogeneity explicitly.

allocation of wealth in savings and stocks and bonds. Further investigations should include housing wealth and clarify this conjecture.

The second panel of the table displays the parameter estimates of the threshold equation. Many of the parameter estimates are significant, including σ_T. This indicates that adding the threshold equation is an improvement compared to the model in which $T^{\star\star} \equiv 0$. This is also indicated by a likelihood ratio test. The results indicate that the threshold for holding financial assets increases with net income but decreases with the marginal tax rate. Together with the corresponding estimate from the wealth equation, the former effect indicates that the higher one's income, the higher is the threshold for investing in financial assets, and given that this threshold is being passed, the higher is the amount invested there. The effect of the tax rate can be caused by preferential taxation especially of stocks and bonds which induces to hold more of them, such that with a rising marginal tax rate the gap between the threshold and the amount of wealth held widens in both directions. We only partly capture arbitraging possibilities which are due to the structure of the Dutch tax system: Interest on savings is tax free up to a level of 1000 Dfl. for singles (and 2000 Dfl. for couples). Beyond that, there is an additional tax allowance of 1000 (2000) Dfl. for dividend income. Thus, including just one threshold does not give full justice to the incentive scheme provided by the tax system. We do, however, control for an influence of the marginal tax rate on the demand for risky assets (see below).

Age has a similar effect on the threshold as it has on wealth, although much more pronounced. The point estimates indicate the threshold function conditional on age to lie below the wealth function for the entire sample. Educational and occupational effects on thresholds could not be detected, by and large. The exceptions are selfemployed individuals who face a higher threshold value. For couples, households with children or those headed by women, the thresholds are found to be lower than for the reference group (single, widowed, divorced, childless). Apparently, these groups are more willing to hold small amounts of financial assets than others. It might reflect that their (subjective) time and money costs are lower. In addition, northern households have significantly lower thresholds than households from the west.

Turning to the third equation, the presumption that the wealthier households hold relatively more risky assets is confirmed by the parameter of the wealth variable log(wealth). Wealth is the most important determinant of the share of stocks and bonds. In terms of consumer demand theory this means that stocks and bonds are a luxury: if financial wealth increases by 1%, the amount invested in stocks and bonds increases by more than 1%. If we rewrite the utility function in terms of returns, it implies that relative risk aversion decreases with wealth: if W increases, the share invested in the risky asset also increases (cf. Arrow (1965) and Pratt (1964)).

The marginal tax rate influences the decision of how to allocate wealth between different assets and induces people to choose more risky assets. This is in line with the predictions from the rather simplistic Domar/Musgrave model on the effects of a proportional income tax with full loss offset, since there the marginal rate of transformation between risk and return of the assets is unaffected and the remaining negative pure wealth effect can only be compensated by taking higher risk. In general however, the theoretical predictions of the effects of taxation on portfolio choice depend on the parameters of the underlying utility function and the distribution of asset returns (Feldstein (1976)) as well as on other features of the tax system, e.g. deductibility of fixed transaction costs (Leape (1987)). Moreover, deductibility of interest payments and exempts in the Dutch tax system, as sketched above, gives rise to the diversification pattern for households with high marginal tax rates observed in these estimates.

The educational level plays a significant role as well. Higher educated people tend to hold more stocks and bonds relative to savings. One might interpret this in favor of the hypothesis of King and Leape (1984, 1987) that asset holdings are determined by the informational status that the investor has acquired with respect to certain assets. Since stocks and bonds can be viewed as information intensive assets and the informational status can be proxied by the educational level of the investor, one could interpret the finding here as supportive of this hypothesis. But — abstracting from cohort effects — there should be an influence of age as well. The quadratic age specification exhibits a strongly significant u–shaped pattern for the relative demand of stocks and bonds, with a minimum at age 46.

Viewing this in the light of liquidity constraints, the result obtained here is somewhat counter–intuitive, since it has been argued that younger people refrain from holding riskier assets due to tighter borrowing restrictions (cf. Paxson (1990) for this effect, if riskier assets are less liquid). If younger workers have a higher labor supply flexibility however, this might insure them against income risk and enable them to hold more risky assets (see Bodie et al. (1992)).

The occupational and professional variables could be interpreted as proxies for income risk, but since they are insignificant we cannot detect such an effect. It should be pointed out, however, that it is quite likely that the occupational or professional choice is subject to a self–selection effect where less risk averse people choose riskier jobs and thus these dummies do not measure the effect of pure income risk (cf. Guiso et al. (1994)).

The estimate for the covariance between the first and third equation in (2) is insignificant however and the correlation coefficient of $\rho_{ws} = 0.099$ does not indicate a pronounced correlation between the level of wealth and the share in risky assets.

In order to assess the fit of the model, we predict the number of observations, the level of the threshold values and the level of the log of wealth for each regime. Four regimes are distinguished, according to whether the amounts invested in stocks and bonds and in saving accounts are positive or zero. For this purpose, we have carried out a simulation from the estimated parameter values. The results are shown in table 6. Comparing the estimated figures with those found in the data shows that the model in its present specification does quite well. Overall, we on average correctly capture 98.9% of the observations in four distinguishable regimes.[11] Only the regimes with positive stocks and bonds but zero savings cannot be recovered accurately, which is due to the small number of observations in this regime in the data. The value of the thresholds below which no wealth in the considered assets is held varies across regimes and is on average substantially higher in the first regime than in the other three. In general, there is not much that can be expected a–priori about the threshold values. By and large, they seem to be of reasonable order. Interestingly, the thresholds are lowest for the case where no stocks and bonds but exclusively saving accounts are held, indicating lower hurdles for the demand for the "riskier" asset.

Since σ_T is small compared to the systematic part of $T^{\star\star}$, the probability of a negative $T^{\star\star}$ will become small. Thus, we can safely ignore the possibility that $T^{\star\star} < W^{\star\star} < 0$ (cf. section II). This is also indicated by the comparison of conditional means to conditional standard deviations per regime, as displayed in table 6.

VI. Conclusions

This paper deals with the question of how people in the Netherlands structure their portfolios in financial wealth. We focus on two major asset categories, namely savings and stocks and bonds, and estimate demand equations in the form of budget shares by an extended tobit model. One of the main determinants of asset demand, namely the overall amount invested in the two asset groups, is taken to be endogenous, and thus the econometric specification explains wealth and the share in risky assets simultaneously.

Compared with other studies in this field, the first innovation provided here is that the econometric model takes non–reporting of endogenous variables fully into account. The usual approach would be to discard all missing values in endogenous variables without exploiting their informational content (Feldstein (1976) even ignores zero–wealth observations). Second, the introduction of a threshold equation improves the fit dramatically In a previous version of this paper we presented the analysis with a deterministic zero threshold for wealth, and discarded for comparison the incomplete observations (about ten percent of the sample). Results, which are available from the authors on request, show that this theoretically incorrect treatment leads to less precise and biased estimates in the shares equation and renders a rather poor fit of the model.

Nevertheless, the model estimated here still has some drawbacks. One is that the specification does not allow for heteroscedastic error terms. This might be overcome by allowing the errors to be a function of wealth, for instance, or by a transformation of the endogenous variable in the shares equation (e.g. as in Ioannides (1992) or Hubbard (1985)). Furthermore, some additional

[11] In the simulations we do not consider the observability problem for positive but non–reported amounts which gives rise to additional 5 regimes (cf. table 1).

scrutiny of the exclusion restrictions in the shares equation might be in order. And in terms of exogenous variables, human capital might play a role in the portfolio choice decision (this is stressed by e.g. Feldstein (1976) or Bodie et al. (1992)). The assumption of trivariate normality can be considered as a further drawback. Semi–parametric estimators for the model at hand, taking account of the unobservability problem, are not yet available, however.

One extension certainly concerns the number of assets chosen. It would be worthwhile to include housing wealth into the portfolio model since the value of the house is the single most important asset in the data. However, the modelling would have to take account of mortgages. From the econometric point of view this would add more regimes to the likelihood and make estimation more cumbersome (we have up to 3^n regimes if n denotes the number of assets.) In this case one would have to consider more closely the non–negativity constraints that in the two asset case can be captured by the tobit specification. This also might lead to different estimation methods (cf. Wales and Woodland (1983)). More structure could be achieved if fixed cost would explicitly be taken into account.

From the theoretical part, it might be worthwhile to leave the framework of the static neoclassical consumer demand model and to look at dynamic models which explain life cycle behavior where agents maximize a discounted expected utility subject to life cycle budget constraints.

References

AGELL, JONAS AND PER-ANDERS EDIN (1990): "Marginal Taxes and the Asset Portfolios of Swedish Households," *Scandinavian Journal of Economics*, 92, 47-64.

ALESSIE, ROB, HERMAN CAMPHUIS AND PRAMILA KRISHNAN (1993a): "Household Wealth and the Composition of Portfolios." Tilburg University, mimeo.

ALESSIE, ROB, MENNO PRADHAN AND CHRISTINE ZANDVLIET (1993b): "An Exploratory Analysis of the Socio–Economic Panel with Regard to the Financial Position of Households." VSB–CentER Savings Project, Progress Report No. 14, Tilburg University.

AMEMIYA, TAKESHI, MAKOTO SAITO AND KEIKO SHIMONO (1993): "A Study of Household Investment Patterns in Japan: An Application of Generalized Tobit Model," *Economic Studies Quarterly*, 44, 13-28.

ARROW, KENNETH J. (1965): *Aspects of the Theory of Risk–bearing.* Helsinki: Yrjö Jahnssonin Säätiö.

BLACKORBY, CHARLES, DANIEL PRIMONT AND R. ROBERT RUSSELL (1978): *Duality, Separability, and Functional Structure: Theory and Economic Applications.* New York: North–Holland.

BANKS, JAMES, RICHARD BLUNDELL AND ARTHUR LEWBEL (1994): "Quadratic Engel Curves, Indirect Tax Reform and Welfare Measurement." Discussion Papers in Economics, 94–04, University College London.

BODIE, ZVI, ROBERT C. MERTON AND WILLIAM F. SAMUELSON (1992): "Labor Supply Flexibility and Portfolio Choice in a Life–cycle Model," *Journal of Economic Dynamics and Control*, 16, 427–449.

CHRISTENSEN, LAURITS R., DALE W. JORGENSON AND LAWRENCE J. LAU (1975): "Transcendental Logarithmic Utility Functions," *American Economic Review*, 65, 367-83.

DEATON, ANGUS S. (1981): "Theoretical and Empirical Approaches to Consumer Demand under Rationing," in *Essays in the Theory and Measurement of Consumer Behaviour,* ed. by Angus Deaton. Cambridge: Cambridge University Press, pp. 55-72.

DEATON, ANGUS S. AND JOHN MUELLBAUER (1980): *Economics and Consumer Behavior.* Cambridge: Cambridge University Press.

ENGEN, ERIC M., WILLIAM G. GALE AND JOHN KARL SCHOLZ (1994): "Do Saving Incentives Work?" *Brookings Papers on Economic Activity*, 1, 85-151.

FELDSTEIN, MARTIN (1976): "Personal Taxation and Portfolio Composition: An Econometric Analysis," *Econometrica*, 44, 631-650.

GOLDSMITH, DAVID (1976): "Transactions Costs and the Theory of Portfolio Selection," *Journal of Finance*, 31, 1127–1139.

GUISO, LUIGI, TULLIO JAPELLI AND DANIELE TERLIZZESE (1994): "Income Risk, Borrowing Constraints and Portfolio Choice." VSB–CentER Savings Project, Progress Report No. 19, Tilburg University.

HUBBARD, GLENN R. (1985): "Personal Taxation, Pension Wealth, and Portfolio Composition," *Review of Economics and Statistics*, 67, 53-60.

IOANNIDES, YANNIS M. (1992): "Dynamics of the Composition of Household Asset Portfolios and the Life Cycle," *Applied Financial Economics*, 2, 145-159.

KING, M. A. AND L.-D. L. DICKS-MIREAUX (1982): "Asset Holdings and the Life Cycle," *The Economic Journal*, 92, 247-267.

KING, MERVYN A. AND JONATHAN I. LEAPE (1984): "Wealth and Portfolio Composition: Theory and Evidence." NBER Working Paper, No. 1468.

KING, MERVYN A. AND JONATHAN I. LEAPE (1987): "Asset Accumulation, Information, and the Life Cycle." NBER Working Paper, No. 2392.

LEAPE, JONATHAN I. (1987): "Taxes and Transaction Costs in Asset Market Equilibrium," *Journal of Public Economics*, 33, 1-20.

MADDALA, G. S. (1983): *Limited–dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

MAYSHAR, JORAM (1981): "Transaction Costs and the Pricing of Assets," *Journal of Finance*, 36, 583-597.

NELSON, FORREST D. (1977): "Censored Regression Models with Unobserved, Stochastic Thresholds," *Journal of Econometrics*, 6, 309-327.

PAXSON, CHRISTINA (1990): "Borrowing Constraints and Portfolio Choice," *Quarterly Journal of Economics*, 105, 535-543.

PHLIPS, LOUIS (1983): *Applied Consumption Analysis.* 2nd ed. Amsterdam: North-Holland.

PRATT, JOHN W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122-136.

SHARPE, WILLIAM F. (1991): "Capital Asset Prices with and without Negative Holdings," *The Journal of Finance*, 46, 489-509.

SHORROCKS, ANTHONY (1982): "The Portfolio Composition of Asset Holdings in the United Kingdom," *The Economic Journal*, 92, 268-284.

UHLER, R. S. AND J. G. CRAGG (1971): "The Structure of the Asset Portfolios of Households," *Review of Economic Studies*, 38, 341-357.

VENTI, STEVEN F. AND DAVID A. WISE (1986): "Tax-Deferred Accounts, Constrained Choice, and Estimation of Individual Saving," *Review of Economic Studies*, 53, 579-601.

WALES, T. J. AND A. D. WOODLAND (1983): "Estimation of Consumer Demand Systems with Binding Non–negativity Constraints," *Journal of Econometrics*, 21, 263-285.

Appendix

This appendix lists the regimes and likelihood contribution for the model (2):

$$\begin{array}{rclclcl} w^{\star\star} & = & x_w\beta_w & + & & & \epsilon_w \\ T^{\star\star} & = & x_T\beta_T & + & & & \epsilon_T \\ s^{\star\star} & = & x_s\beta_s & + & w^\star\gamma & + & \epsilon_s \end{array}$$

non–negativity constraints:

$$w^\star = w^{\star\star}\mathbf{1}_{[w^{\star\star}>T^{\star\star}]}$$

$$s^\star = \begin{cases} 0 & \text{if} \quad s^{\star\star} \leq 0 \\ s^{\star\star} & \text{if} \quad 0 < s^{\star\star} < 1 \\ 1 & \text{if} \quad s^{\star\star} \geq 1 \end{cases}$$

Assuming normality with

$$\Sigma = \begin{pmatrix} \sigma_w^2 & \rho_{wT}\sigma_w\sigma_T & \rho_{ws}\sigma_w\sigma_s \\ \cdot & \sigma_T^2 & \rho_{Ts}\sigma_T\sigma_s \\ \cdot & \cdot & \sigma_s^2 \end{pmatrix}$$

plus the restrictions $\rho_{Ts} = \rho_{wT} = 0$ (i.e. independence under normality).

In the following W denotes total wealth and $w = \log W$. For the reduced form contributions (regimes 3, 5, and 9), denote the errors $v = \epsilon_s + \gamma\epsilon_w$, $u = \epsilon_w - \epsilon_T$ with

$$\begin{array}{rcl} \sigma_v & = & \sqrt{\sigma_s^2 + \gamma^2\sigma_w^2 + 2\gamma\sigma_{sw}} \\ \sigma_u & = & \sqrt{\sigma_T^2 + \sigma_w^2} \\ \rho_{uv} & = & \dfrac{\sigma_{sw} + \gamma\sigma_w^2}{\sigma_v\sigma_u} \end{array}$$

Univariate pdf, cdf and bivariate cdf of the standard normal are denoted by $\phi(\cdot)$, $\Phi(\cdot)$, $\Phi_{\text{bv}}(\cdot,\cdot,\cdot)$, respectively.
Likelihood contributions $\mathcal{L}$ (regimes coded as in table 1):

1. $W > 0$:
$$\mathcal{L} = \Pr(w^{\star\star} \leq T^{\star\star}) = \Phi\left(\frac{x_T\beta_T - x_w\beta_w}{\sigma_u}\right)$$

2. $W > 0$, observed, $s = 0$:
$$\mathcal{L} = \Pr(s^{\star\star} \leq 0, T^{\star\star} < w | w^{\star\star} = w) \cdot f_{w^{\star\star}}(w)$$
with
$$f_{w^{\star\star}}(w) = \frac{1}{\sigma_w}\phi\left(\frac{w - x_w\beta_w}{\sigma_w}\right)$$
and
$$\begin{aligned} &\Pr(s^{\star\star} \leq 0, T^{\star\star} < w | w^{\star\star} = w) = \\ &= \Phi\left(\frac{-x_s\beta_s - w\gamma - \rho_{sw}\sigma_s/\sigma_w\,(w - x_w\beta_w)}{\sigma_s\sqrt{1-\rho_{sw}^2}}\right) \cdot \Phi\left(\frac{w - x_T\beta_T}{\sigma_T}\right) \end{aligned}$$

3. $W > 0$, unobserved, $s = 0$:

$$\begin{aligned}\mathcal{L} &= \Pr(s^{\star\star} \leq 0, w^{\star\star} > T^{\star\star}) = \\ &= \Pr(v \leq -[x_s\beta_s + x_w\beta_w\gamma], u > x_T\beta_T - x_w\beta_w) = \\ &= \Phi_{\text{bv}}\Big(\frac{-[x_s\beta_s + x_w\beta_w\gamma]}{\sigma_v}, \frac{x_w\beta_w - x_T\beta_T}{\sigma_u}, -\rho_{uv}\Big)\end{aligned}$$

Since $w^\star$ is unobserved in this case, and $w^{\star\star} = w^\star$, the reduced form has to be exploited.

4. $W > 0$, observed, $s = 1$ (analogous to regime 2):

$$\begin{aligned}\mathcal{L} &= \Pr(s^{\star\star} > 1, T^{\star\star} < w | w^{\star\star} = w) \cdot f_{w^{\star\star}}(w) = \\ &= \Big[1 - \Phi\Big(\frac{1 - x_s\beta_s - w\gamma - \rho_{sw}\sigma_s/\sigma_w\,(w - x_w\beta_w)}{\sigma_s\sqrt{1-\rho_{sw}^2}}\Big)\Big] \cdot \Phi\Big(\frac{w - x_T\beta_T}{\sigma_T}\Big) \\ & \quad \frac{1}{\sigma_w}\phi\Big(\frac{w - x_w\beta_w}{\sigma_w}\Big)\end{aligned}$$

5. $W > 0$, unobserved, $s = 1$ (analogous to regime 3):

$$\mathcal{L} = \Pr(s^{\star\star} \geq 1, w^{\star\star} > T^{\star\star}) = \Phi_{\text{bv}}\Big(\frac{-1 + x_s\beta_s + x_w\beta_w\gamma}{\sigma_v}, \frac{x_w\beta_w - x_T\beta_T}{\sigma_u}, \rho_{uv}\Big)$$

6. $W > 0, 0 < s < 1$, both observed:

$$\begin{aligned}\mathcal{L} &= f_{s^{\star\star},w^{\star\star}}(s,w) \cdot \Pr(T^{\star\star} < w | w, s) = \\ &= \frac{1}{2\pi\sigma_w\sigma_s\sqrt{1-\rho_{sw}^2}} \cdot \\ &\cdot \exp\Big\{-\frac{(w - x_w\beta_w)^2\sigma_s^2 - 2(w - x_w\beta_w)(s - w\gamma - x_s\beta_s)\sigma_{sw} + (s - w\gamma - x_s\beta_s)^2\sigma_w^2}{2\sigma_w^2\sigma_s^2(1-\rho_{sw}^2)}\Big\} \cdot \\ &\cdot \Phi\Big(\frac{w - x_T\beta_T}{\sigma_T}\Big)\end{aligned}$$

7. Let a_1 denote the amount of stocks and bonds held;
$W > 0, \;\; 0 < s < 1, a_1 = sW$ observed, $(1-s)W$ not observed:

$$\mathcal{L} = \int_0^1 f_{a_1^\star, s^{\star\star}}(a_1, s^{\star\star}) \Pr(T^{\star\star} > w^{\star\star} | a_1, s^{\star\star}) \mathrm{d}s^{\star\star}$$

with

$$\Pr(T^{\star\star} > w^{\star\star} | a_1, s^{\star\star}) = \Phi\Big(\frac{\ln(a_1/s^{\star\star}) - x_T\beta_T}{\sigma_T}\Big)$$

$f(\cdot,\cdot)$ is the density of $(a_1^\star, s^{\star\star}) = (s^{\star\star}\exp(w^{\star\star}), s^{\star\star})$, obtained from the bivariate normal density of $(w^{\star\star}, s^{\star\star})$ and including a Jacobian. Numerical integration is required.

8. Let a_2 denote the amount of savings held;
$W > 0, \;\; 0 < s < 1, a_2 = (1-s)W$ observed, sW not observed:

$$\mathcal{L} = \int_0^1 f_{a_2^\star, s^{\star\star}}(a_2, s^{\star\star}) \Pr(T^{\star\star} > w^{\star\star} | a_2, s^{\star\star}) \mathrm{d}s^{\star\star}$$

with

$$\Pr(T^{\star\star} > w^{\star\star} | a_2, s^{\star\star}) = \Phi\Big(\frac{\ln(a_2/(1 - s^{\star\star})) - x_T\beta_T}{\sigma_T}\Big)$$

9. $W > 0,\;\; 0 < s < 1$, both stocks and bonds and savings are unobserved:

$$\begin{aligned}\mathcal{L} &= \Pr(0 < s^{\star\star} < 1, w^{\star\star} > T^{\star\star}) = \\ &= \Phi_{\mathrm{bv}}\Big(\frac{x_s\beta_s + x_w\beta_w\gamma}{\sigma_v}, \frac{x_w\beta_w - x_T\beta_T}{\sigma_u}, \rho_{uv}\Big) - \Phi_{\mathrm{bv}}\Big(\frac{-1 + x_s\beta_s + x_w\beta_w\gamma}{\sigma_v}, \frac{x_w\beta_w - x_T\beta_T}{\sigma_u}, \rho_{uv}\Big)\end{aligned}$$

Tables

Table 1: number of observations per regime

number of observations (%)	savings = 0	savings > 0 (observed)	savings > 0 (unobserved)	sum
stocks & bonds = 0	regime 1 513 (16.18)	regime 2 2192 (69.13)	regime 3 266 (8.39)	2971 (93.69)
stocks & bonds > 0 (observed)	regime 4 8 (0.25)	regime 6 114 (3.59)	regime 7 39 (1.23)	161 (5.08)
stocks & bonds > 0 (unobserved)	regime 5 1 (0.03)	regime 8 28 (0.88)	regime 9 10 (0.32)	39 (1.23)
sum	522 (16.46)	2334 (73.60)	315 (9.93)	3171 (100.00)

Table 2: Summary statistics of variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
stocks & bonds	3132	3765.3	39776	0	974000
savings	2856	13341	50493	0	983095
wealth	2827	16512	69984	0	1415249
fraction of stocks & bonds	2314	0.0273	0.1367	0	1
log(income+1)	3171	7.7361	0.7639	0	12.92
income	3171	3535.2	14062	0	410000
marginal tax rate	3171	0.4837	0.0720	0	0.7
age of head	3171	44.004	15.422	18	89
age squared / 1000	3171	2.1741	1.5165	0.324	7.921
high education	3171	0.1583	0.3651	0	1
high education * age	3171	6.4566	15.667	0	77
full-time work	3171	0.5976	0.4905	0	1
part-time work	3171	0.0489	0.2157	0	1
farmer	3171	0.0076	0.0867	0	1
selfemployed	3171	0.0952	0.2936	0	1
whitecollar	3171	0.4447	0.4970	0	1
bluecollar	3171	0.2491	0.4326	0	1
married or living together	3171	0.6976	0.4594	0	1
female	3171	0.2021	0.4017	0	1
children	3171	0.979	1.1221	0	8
small children	3171	0.2983	0.6545	0	4
region: north	3171	0.3409	0.4741	0	1
region: south	3171	0.1962	0.3971	0	1

Definition of variables:

`savings`: sum of the amounts held in saving account balances, time deposit accounts and saving certificates and certificates of deposit.

`stocks and bonds`: shares in domestic companies, shares in foreign companies, shares in investment fonds, options, bonds and mortgage bonds.

`wealth`: sum of savings and stocks and bonds

`fraction of stocks and bonds`: is the share of this asset category in `wealth`

`income`: the sum of monthly net income (in Dfl) of the head of the household and his or her partner. 7 households report zero income. The income sources are not specified. Since it is monthly, income probably excludes asset income.

`marginal tax rate`: calculated from individual net income based on the parameters of the Dutch tax system in 1988; the household tax rate is chosen to equal the maximum of the individual tax rates within the household.

`Age`: measured in years.

The other explanatory variables are dummy variables (except for the number of `children` and small children). The educational status is captured by the variable on high education (university degree or higher vocational school), the reference group includes everything else.

Labor supply is captured by the indicators for `full-time` (36 hours per week or more) and `part-time` employment (10 – 35 hours), the reference group consists of people who work less or not at all (disabled, unemployed, retired, students and housewives/men without alternative occupation).

Professional status is described by the variables `farmer` (farmer or market gardener), `selfemployed` (includes free lancers, managers or owners of firms), `whitecollar` employees, and `bluecollar` workers. Reference group here are people without paid employment and others.

The variable `small children` gives the number of children of age 6 and below.

The remaining variables on family composition and marital status are self–explaining. Regional dummy variables distinguish between `north` (provinces Groningen, Friesland, Drenthe, Overijssel, Gelderland) and `south` (Noord Brabant and Limburg) of the country (reference group: Noord and Zuid Holland, Utrecht, Zeeland, Flevoland).

Table 3: structure of diversification in 4 major asset categories

	stocks & bonds	saving accounts	life insurance	housing wealth	Obs.	Percent
1	yes	yes	yes	yes	89	2.81
2	yes	yes	yes	no	12	0.38
3	yes	yes	no	yes	65	2.05
4	yes	yes	no	no	25	0.79
5	yes	no	yes	yes	1	0.03
6	yes	no	yes	no	4	0.13
7	yes	no	no	yes	4	0.13
8	yes	no	no	no	0	0.00
9	no	yes	yes	yes	710	22.39
10	no	yes	yes	no	436	13.75
11	no	yes	no	yes	544	17.16
12	no	yes	no	no	768	24.22
13	no	no	yes	yes	65	2.05
14	no	no	yes	no	101	3.19
15	no	no	no	yes	72	2.27
16	no	no	no	no	275	8.67
	total				3171	100.00

Table 4: level of diversification in 4 major asset categories

number of asset categories	Obs.	Percent	Cumulative
0	275	8.67	8.67
1	941	29.68	38.35
2	1078	34.00	72.34
3	788	24.85	97.19
4	89	2.81	100.00
Total	3171	100.00	

Table 5: Estimation results

dependent variable:	log of wealth $w^{\star\star}$			threshold $T^{\star\star}$			fraction of stocks and bonds $s^{\star\star}$		
Variable	estimate	stddev.	t-value	estimate	stddev.	t-value	estimate	stddev.	t-value
constant	1.953	0.493	3.958	3.601	0.763	4.718	-4.097	0.651	-6.294
log(income+1)	0.471	0.053	8.962	0.536	0.099	5.421			
marginal tax rate	2.064	0.790	2.612	-3.060	1.398	-2.188	2.538	0.810	3.133
age of head	-0.007	0.020	-0.340	-0.047	0.027	-1.725	-0.036	0.017	-2.150
age squared / 1000	0.445	0.208	2.145	0.877	0.281	3.122	0.391	0.183	2.132
high education	-0.639	0.369	-1.730	-0.062	0.597	-0.103	0.350	0.093	3.782
age * high education	0.023	0.009	2.603	0.006	0.015	0.396			
full-time work	0.723	0.131	5.504	0.182	0.198	0.923	0.063	0.175	0.358
part-time work	0.140	0.197	0.713	-0.170	0.289	-0.588	0.146	0.203	0.716
farmer	1.197	0.496	2.411	-0.746	1.142	-0.653			
selfemployed	0.048	0.196	0.243	1.017	0.296	3.437	0.135	0.171	0.794
white-collar	-0.182	0.138	-1.322	-0.274	0.199	-1.373	0.155	0.145	1.071
blue-collar	-0.537	0.145	-3.691	0.193	0.198	0.973	-0.267	0.184	-1.448
married or lvng. tog.	0.368	0.141	2.605	-0.678	0.196	-3.464			
female	-0.251	0.144	-1.741	-0.707	0.200	-3.536			
number of children	0.010	0.051	0.199	-0.273	0.079	-3.440			
small children	-0.094	0.077	-1.217	-0.188	0.135	-1.390			
region: north	0.184	0.090	2.053	-0.350	0.134	-2.613			
region: south	0.196	0.115	1.695	-0.114	0.172	-0.660	-0.155	0.103	-1.514
log(wealth)							0.240	0.094	2.560
covariance matrix of error terms									
stddev wealth				2.008	0.045	44.667			
stddev thresh.				1.352	0.044	30.394			
stddev shares				0.795	0.069	11.575			
corr. wealth/thresh.				0.0	fixed				
corr. wealth/shares				0.099	0.219	0.452			
corr. shares/thresh.				0.0	fixed				

Log-likelihood:	-26520.1502
number of obs.:	3171

Table 6: model performance[a]

	number of observations		log wealth (held)		simulated threshold values	
	data	simul.	data	simul.	mean	stddev.
stocks & bonds = 0 savings = 0	513	501	0.00	0.00	6.59	1.50
stocks & bonds = 0 savings $>$ 0	2458	2456	8.16	8.15	4.90	1.55
stocks & bonds $>$ 0 savings = 0	9	27	10.69	11.74	5.63	1.74
stocks & bonds $>$ 0 savings $>$ 0	191	187	10.41	10.46	5.36	1.71

[a]the actual values are for regimes 1, 2, 4, and 6, the predicted values include information from regimes 1, 2 & 3, 4 & 5, and 6 – 9, respectively; the expected values for log wealth and the thresholds have been calculated by simulation, based on 50 random draws