Preliminary

SAVING AND REFERENCE GROUPS

Arie Kapteyn

CentER

Tilburg University

January 2000

Abstract

I consider how the saving behavior of consumers is affected by incomes in their reference group. Although standard life cycle models with interdependent preferences provide no clear cut prediction of how incomes in one's reference group should influence savings, the empirical analysis points to an unambiguously negative effect: If incomes in the reference group are higher, savings are lower. The data used stem from the CentER savings survey, in which households are asked direct questions about both their reference groups and their savings behavior.

1. Introduction

For a very long time the study of interdependent preferences has remained at the fringe of the economics profession. Although the notion that individual preferences are influenced by the behavior of others has been around for a very long time, the number of empirical studies taking such influences into account is stil quite modest (See Kapteyn et al. (1997) and Brock and Durlauf (forthcoming) for references to some exceptions). One can think of several reasons why interdependence of preferences is so rarely studied. Perhaps the most fundamental one is just lack of adequate data. As was pointed out by Manski (1993), if a researcher has no independent information about the reference group of individuals (i.e. the group of people whose behavior may influence the preferences of an individual), reference group effects are generally not identified. Since such information is indeed rarely available, studies purporting to measure reference group influences have to make (more or less plausible) assumptions about the exact nature of reference goups (as was done for instance by Van de Stadt et al. (1985) or Kapteyn et al. (1997). An additional problem is that even if we know what people's reference groups are, conformity in behavior may result, for instance, from similarity in characteristics (denoted as *correlated effects* by Manski) or from the fact that people truly influence each other (*endogenous effects)*. In addition there is the possibility of *exogenous effects* (in Manski's terminology): for instance the availability of certain consumer goods may depend on the socio-economic characteristics of a group (e.g., fancy shops only locate in wealthy neighborhoods) and thereby influence consumption behavior. The identification problem, of how to distinguish endogenous effects from other possibilities, has become the main stumbling block, not just in the empirical modelling of interdependent preferences, but also in related areas like the analysis of neighborhood effects, diffusion of innovations, etc. (cf. Brock and Durlauf, forthcoming).

In an earlier study (Woittiez and Kapteyn (1999)), we have used direct questions to respondents about some characteristics of people in their reference group to study the labor supply decisions of married women. The basic idea in that paper has been to use information about the characteristics of people in the reference group of an individual to construct measures of average labor supply in the reference group of that individual. The average labor supply is next used as an explanatory variable for the observed labor supply of the individual herself. The problem with that approach is that we cannot be sure that it is actually the constructed reference group variable itself (for instance: average hours worked by females in the reference group of the respondent) which explains observed behavior, or whether the characteristics used to construct the reference group variables (e.g. education or age in the reference group) have a direct influence on her labor supply (this might represent

exogenous effects). And again: if we observe a relationship between the hours worked in one's reference group and own hours, it may still only signify similarity of background characteristics as a result of sorting (correlated effects).

In the current paper, I am considering reference group effects in saving. Although I am not able to avoid all potential pitfalls, the empirical analysis presents an improvement over earlier work in a number of ways. First of all, I do not relate savings by individuals to savings by others, but rather savings by individuals to income of others. Since also own income will be used as an explanatory variable, this would seem to avoid, at least to an extent, exogenous effects as a possible explanation of effects that we find. Correlated effects will be harder to exclude, since there is still the possibility that people who save a lot (or not at all) tend to associate mainly with others who have particularly low (or high) incomes.

To the best of my knowledge, no empirical economic studies of reference group effects on saving behavior exist to-day. A paper that comes close in spirit is Binder and Pesaran (forthcoming), but their paper is primarily theoretical.

Another reason why interdepence is seldomly incorporated in empirical models, may be that behavioral implications are ambiguous: To understand how reference groups may affect savings behavior, I discuss a simple two-period model of interdependent preferences in Section 2. I investigate under what circumstances, one may expect a clear-cut effect of reference group incomes on the savings of individuals. It is worth pointing out that in many cases one *cannot* discern any observable effect of reference group variables on individual saving behavior. The implication of this is, that preference interdependence may exist, and yet have no implications for observable behavior. Section 3 describes the data used in the empirical analysis. A preliminary analysis suggests the existence of a negative relation between reference group incomes and savings. Section 4 expands the empirical analysis, without changing this conclusion. Section 5 provides some concluding remarks.

Appendix A presents the wording of some of the main survey questions used in the empirical analysis.

2. Examples in a two-period framework

To motivate the empirical analysis, I restrict myself to a simple two-period two-agent framework, with complete certainty, quadratic preferences and no habit formation. For most of the points to be made this simple framework will suffice, although Appendix B[1] presents a generalization incorporating habit formation.

Thus, consider two agents with utility functions:

$$\begin{aligned} U_1 &= -\tfrac{1}{2}(c_1-\alpha_1)^2 - \tfrac{1}{2}\delta(c_2-\alpha_2)^2 \\ U_2 &= -\tfrac{1}{2}(d_1-\beta_1)^2 - \tfrac{1}{2}\phi(d_2-\beta_2)^2 \end{aligned} \qquad \textbf{(1)}$$

where c_τ and d_τ (τ=1, 2) are the consumption levels in consecutive periods of agents 1 and 2 respectively; α_τ and β_τ (τ=1, 2) are parameters of the instantaneous utiltity functions; δ and ϕ represent time preference rates (higher values indicate *less* discounting of the future)[2]. Let the endowments of the agents 1 and 2 at the beginning of period 1 be Y and Z respectively. Maximizing the utility functions subject to $c_1 + c_2 = Y$ and $d_1 + d_2 = Z$ respectively, yields:

$$\begin{aligned} c_1 &= \alpha_1 + \frac{\delta}{1+\delta}(Y-\alpha_1-\alpha_2) \\ c_2 &= \alpha_2 + \frac{1}{1+\delta}(Y-\alpha_1-\alpha_2) \\ d_1 &= \beta_1 + \frac{\phi}{1+\phi}(Z-\beta_1-\beta_2) \\ d_2 &= \beta_2 + \frac{1}{1+\phi}(Z-\beta_1-\beta_2) \end{aligned} \qquad \textbf{(2)}$$

To represent interdependence of preferences, we parameterize the instantaneous utility parameters as follows:

---

[1]Appendix B is still missing in this version of the paper.

[2]Without loss of generality, I have taken the interest rate to be equal to zero. Hence, the time preference rates can be interpreted as the ratio of a time preference rate and the interest rate.

$$a \equiv \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \\ \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \psi_1 d_1 \\ \psi_2 d_2 \\ \gamma_1 c_1 \\ \gamma_2 c_2 \end{pmatrix} \equiv a_0 + \begin{pmatrix} \psi_1 d_1 \\ \psi_2 d_2 \\ \gamma_1 c_1 \\ \gamma_2 c_2 \end{pmatrix} \tag{3}$$

I will assume $0<\psi_1<1$, $0<\psi_2<1$, $0<\gamma_1<1$, $0<\gamma_2<1$, i.e. envy rather than altruism, as I consider that to be the empirically more relevant case.

To find equilibrium values of consumption we have to solve (2) and (3) for c_1, c_2, d_1 and d_2. Define ι_2 as a vector of two unit elements, $E_2 \equiv \iota_2 \iota_2'$, and I_2 as the identity matrix of order 2. Furthermore define:

$$\Delta_1 \equiv \frac{\delta}{1+\delta};\ \Delta_2 \equiv \frac{1}{1+\delta};\ \Delta_3 \equiv \frac{\phi}{1+\phi};\ \Delta_4 \equiv \frac{1}{1+\phi}\ ,\ \Delta \equiv diag(\Delta_1,\Delta_2,\Delta_3,\Delta_4)$$

$$\Gamma_1 \equiv \begin{pmatrix} \psi_1 & 0 \\ 0 & \psi_2 \end{pmatrix};\ \Gamma_2 \equiv \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix};\ \Gamma \equiv \begin{pmatrix} 0 & \Gamma_1 \\ \Gamma_2 & 0 \end{pmatrix};\ y \equiv \begin{pmatrix} Y \\ Z \end{pmatrix};\ \bar{c} \equiv \begin{pmatrix} c_1 \\ c_2 \\ d_1 \\ d_2 \end{pmatrix} \tag{4}$$

Then we can write (2) as

$$\bar{c} = a + \Delta(y \otimes \iota_2) - \Delta(I_2 \otimes E_2)a \tag{5}$$

Notice that (3) can be written as

$$a = a_0 + \Gamma\bar{c} \tag{6}$$

Inserting (6) into (5) yields

$$\bar{c} = a_0 + \Gamma\bar{c} + \Delta(y \otimes \iota_2) - \Delta(I_2 \otimes E_2)a_0 - \Delta(I_2 \otimes E_2)\Gamma\bar{c} \tag{7}$$

Or,

$$\left[I-\Gamma+\Delta(I_2\otimes E_2)\Gamma\right]\bar{c}=a_0+\Delta(y\otimes\iota_2)-\Delta(I_2\otimes E_2)a_0 \tag{8}$$

Using the fact that $\Delta_1+\Delta_2=1$ and $\Delta_3+\Delta_4=1$, one can easily verify that

$$\left[I-\Gamma+\Delta(I_2\otimes E_2)\Gamma\right]=\begin{pmatrix} I_2 & \begin{pmatrix}1\\-1\end{pmatrix}(-\psi_1\Delta_2 \quad \psi_2\Delta_1) \\ \begin{pmatrix}1\\-1\end{pmatrix}(-\gamma_1\Delta_4 \quad \gamma_2\Delta_3) & I_2 \end{pmatrix} \tag{9}$$

Define

$$\Phi_1\equiv\psi_1\Delta_2+\psi_2\Delta_1 \;;\; \Phi_2\equiv\gamma_1\Delta_4+\gamma_2\Delta_3 \;;\; \Phi\equiv\Phi_1\Phi_2 \tag{10}$$

It follows from the assumptions on ψ_1, ψ_2, γ_1, γ_2 that $\Phi<1$. One can verify by direct multiplication that

$$\left[I-\Gamma+\Delta(I_2\otimes E_2)\Gamma\right]^{-1}=$$

$$I+\frac{1}{1-\Phi}\begin{pmatrix} \Phi_1\begin{pmatrix}1\\-1\end{pmatrix}(\gamma_1\Delta_4 \quad -\gamma_2\Delta_3) & \begin{pmatrix}1\\-1\end{pmatrix}(\psi_1\Delta_2 \quad -\psi_2\Delta_1) \\ \begin{pmatrix}1\\-1\end{pmatrix}(\gamma_1\Delta_4 \quad -\gamma_2\Delta_3) & \Phi_2\begin{pmatrix}1\\-1\end{pmatrix}(\psi_1\Delta_2 \quad -\psi_2\Delta_1) \end{pmatrix} \tag{11}$$

Thus, it follows from (8) that

$$\bar{c} = a_0 + \Delta(y \otimes \iota_2) - \Delta(I_2 \otimes E_2)a_0 +$$

$$\frac{1}{1-\Phi}\begin{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}[\Phi(\alpha_{10}\Delta_2 - \alpha_{20}\Delta_1) + \Phi_1(\beta_{10}\Delta_4 - \beta_{20}\Delta_3)] \\ \begin{pmatrix} 1 \\ -1 \end{pmatrix}[\Phi_2(\alpha_{10}\Delta_2 - \alpha_{20}\Delta_1) + \Phi(\beta_{10}\Delta_4 - \beta_{20}\Delta_3)] \end{pmatrix} + \tag{12}$$

$$\frac{1}{1-\Phi}\begin{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}[\Phi_1(\gamma_1\Delta_4\Delta_1 - \gamma_2\Delta_3\Delta_2)Y + (\psi_1\Delta_2\Delta_3 - \psi_2\Delta_1\Delta_4)Z] \\ \begin{pmatrix} 1 \\ -1 \end{pmatrix}[(\gamma_1\Delta_4\Delta_1 - \gamma_2\Delta_3\Delta_2)Y + \Phi_2(\psi_1\Delta_2\Delta_3 - \psi_2\Delta_1\Delta_4)Z] \end{pmatrix}$$

The first row of equation (12) represents consumption behavior without interdependence. The second row represents the effect of interdependence on the intercept of the consumption equations. Generally, I will ignore the intercept terms, as these are not affected by changes in Y or Z. The third row represents the effects of changes in the lifetime resources of the agents on consumption of both agents through the interdependence effects. It is instructive to look at a number of special cases.

Case 1: $\psi_1 = \psi_2$. That is, the influence of consumption of agent 2 on the preferences of agent 1 is the same in both periods. An increase of Z increases consumption by agent 1 in the first period (and hence reduces consumption in the second period) if $\Delta_2\Delta_3 > \Delta_1\Delta_4$. In view of the definitions of the Δs, this is equivalent to $\phi > \delta$. In other words, an increase in resources of agent 2 will increase consumption of agent 1 in the first period if he is less patient than agent 2. Interestingly enough, the interdependence effect in turn causes the consumption of agent 2 to go up in the first period as well (beyond the increase implied by the first row of (12). The reason for this is simple. If agent 1 increases consumption in period 1, this increases the marginal utility of consumption by agent 2 in the same period and he therefore increases his consumption in the first period as well. In a sense the impatience of agent 1 induces impatience in agent 2. A similar observation was made by Binder and Pesaran (forthcoming).

Case 2: $\psi_1 = \psi_2$ and $\phi = \delta$: In this case $\Delta_1 = \Delta_3$ and $\Delta_2 = \Delta_4$. It follows immediately from (12) that changes in Z have no effect on the consumption of agent 1. This was to be expected in view of the discussion of the previous case, as time preference rates are now equal. An increase in Z lowers agent 1's utility, but he has no way of improving his position by shifting consumption between periods. There is still an interdependence effect left, since the response of agent 1 to an increase in his own endowment is different from the case where there is no interdependence. If there holds in addition: $\gamma_1 = \gamma_2$, there is no effect of changes in the endowments of one

agent on the consumption of the other one.

Thus in this framework, one needs to allow for heterogeneity of parameters in order to motivate interdepence effects on consumption, as was pointed out earlier by Binder and Pesaran (1999). Moreover, the sign of the interdependence effects are hard to predict. In a sense one has to know whether agents are in the first period or the second period (in the first part or the second part of one's life) and who is impatient and who is not. An interesting case arises in a simple overlapping generations case. Suppose that the agents only influence each other in period 1 and not in period 2. This might arise if agent 1 first lives in period 2 and then in period 1 (the time order of periods is immaterial in the model), whereas agent 2 first lives in period 1 and then in period 2. This case is captured by:

Case 3: $\psi_2 = \gamma_2 = 0$. Equation (12) shows immediately that an increase in Z unambiguously raises consumption of agent 1 in period 1 (and *vice versa* an increase in *Y* increases consumption by agent 2 in period 1.). Interestingly enough, this non-cooperative equilibrium does not maximize utility of the agents. They would be better off to reduce consumption in period 1 and to consume most of their endowment in the period where they do not influence each other. For example, if we consider the following parameter values: $\alpha_{10}=\alpha_{20}=\beta_{10}=\beta_{20}=4$; $\psi_1=\gamma_1=.5$; $\psi_2=\gamma_2=0$; $\delta=\Phi=.9$; $Y=Z=2$. We find that the equilibrium values for first period consumption of both agents are equal to 1.57, with associated utility equal to -10.9. If there were no interdependence (i.e. $\psi_1=\gamma_1=\psi_2=\gamma_2=0$), the optimal first period consumption levels for both agents would be 1.16. If the agents would stick to this consumption level in the first period in the case where $\psi_1=\gamma_1=.5$ (and still $\psi_2=\gamma_2=0$) their utility level would be -10.3 an improvement over the equilibrium values by 0.6. They could do even better than that: if they choose first period consumption equal to 0.5, for instance, their utility levels would rise to -9.8, i.e. an improvement over the equilibrium values by 1.1.

An alternative interpretation of case 3 is that consumption in period 2 is in the form of a bequest. If the weight given by agent 1 to a bequest is lower than the weight given to current consumption, an increase in the endowment of agent 2 would lead to higher current consumption and hence to a lower bequest. In empirical analysis this would be measured as a lower saving rate of agents with "rich" reference groups.

Case 4: $\psi_1>\psi_2$, $\gamma_1>\gamma_2$. The interpretation of this case is akin to hyperbolic discounting (cf. e.g. Laibson (1998) in the sense that one believes to be less susceptible to the consumption behavior of others in the future than now. For simplicity, assume that $\phi=\delta$. In that case (12) implies that a increase in the endowment agent 1 will increase consumption of agent 2 in the period 1 and *vice versa*.

3. The data and preliminary analysis.

3.1 The data

In the empirical analysis I wil use data from the so-called *CentER Savings Survey* (CSS)[3]. The CSS contains data that are collected annually from the members of the so-called *CentERpanel*. The CentERpanel comprises some 2000 households in the Netherlands. The members of those households fill in a questionnaire at their home computers every week. The CentERpanel is representative of the Dutch population. The questionnaires filled out by the respondents include annual questions about wealth holdings and income. A rather unique aspect of the data is that subjective questions are asked about a wide array of topics, including expectations, risk aversion, time preference, investment strategies, composition of reference groups, etc. The panel started in 1993. I will be using data from the period 1994-1998. Although about half of the respondents receive a computer from the survey organization (CentERdata), which is potentially a powerful incentive to participate in the panel, the heavy response burden (up to 25 minutes per week) leads to a fairly high attrition rate. Thus the panel data I will be using is highly unbalanced, and probably selective.

In this section I present preliminary results mainly based on the so-called *psychological module*, which contains most of the subjective questions mentioned above. The most relevant questions have been cited in Appendix A. It should be noted that in many cases questions are only asked to household heads and that in some cases respondents have the option "don't know" whereas in other cases they don't. As a result, the number of observations may vary quite a bit across variables.

For the purpose of operationalizing reference group influences on consumption, the question KENINK, which asks for an estimate of after tax household incomes in one's circle of acquaintances (see Appendix A) is of particular importance. The frequency distribution of the answers is given in Table 1. For comparison I also give the frequency distribution of self-reported household income in Table 2 (cf. INKHH in Appendix A). A comparison of the distribution of both incomes is slightly hampered by the fact that the brackets for both income measures are different. The distribution of reference group incomes appears to be lying somewhat to the left of the respondents' own income. To facilitate comparison, I have combined income brackets of the reference group incomes so that they coincide with those of the household incomes. Table 3 presents a cross tabulation for both income measures, now using identical brackets. The marginals confirm the observation above that reported reference group incomes tend to be somewhat lower than own household incomes. Not

[3]One can download the data from http://center-ar.kub.nl/website.php3?p=cssindex&l=0.

surprisingly there is a fairly strong positive association between own household income and reference group income; the rank correlation equals 0.61.

Table 4 provides a cross tabulation of own family size and the reported family size in the respondents' reference groups (see KENHH in Appendix A). Again there is a fairly strong positive association between own family size and reference group family size; the rank correlation equals 0.52

We consider two self-reported saving measures. The first one is a combination of OPZIJ (a binary variable indicating whether one put aside any money in the past 12 months) and HOEVOPZIJ (the amount of money put aside if OPZIJ="yes", reported in brackets). The resulting variable SAVING1 ranges from 0 through 7. A second saving measure, SAVING2, is a trinary variable constructed on the basis of INKEVEN. SAVING2 equals -1 if over the past 12 months expenditures were reported to be higher than income; it equals 0 if expenditures were about equal to income; it equals 1 if expenditures were lower than income. Table 5 presents the frequency distribution of SAVING1. Table 6 presents a cross tabulation between SAVING2 and OPZIJ. Although clearly the number of respondents reporting to have put money aside is much greater when ependitures are reported to have been below income than otherwise, more than half of those with expenditures higher than income still claim to have put money aside. This suggests a less than full agreement among respondents of what it means to have put money aside. Table 7 provides a cross tabulation of both saving measures. The rank correlation between both measures is a modest .45.

Table 8 provides the age distribution of the reference groups (See KENLTD in Appendix A) and compares it to the age distribution of the respondents. It appears that generally respondents refer relatively more often to people who are somewhat younger than they are themselves. Since a cross tabulation becomes somewhat unwieldy, I do not present it here. The rank correlation between own age and age in the reference group is quite strong: 0.81 ($\chi^2(132)=46663$).

Table 9 gives a cross tab of the educational attainment of the respondents and of people in their reference groups. Not surprisingly both are correlated (the rank correlation is 0.58), but the reference group education levels reported by the respondents are clearly less dispersed than the education levels of the respondents themselves. The phrasing of the question (see KENOPL in Appendix A) may be causing this, since essentially respondents are asked to give the mode of the educational distribution of their reference group.

The main purpose of the paper is to eplore the extent to which (selfreported) saving behavior of individuals can be explained by reference group variables. In explaining saving behavior we want to control for other factors as much as possible. One of these factors is fluctuations in income. The psychological module contains a question, INKVER, about whether income over the past 12 months increased (recoded in the empirical analysis as 1), whether it stayed about the samel (recoded as 0) or whether it fell (recoded as -1).

Table 10 presents the frequency distribution of this variable. In the sequel the recoded variable will be denoted as INCRISE.

As a final indication of the composition of one' reference group, Table 11 shows the frequency distribution of the answers to the question "What kind of employmennt do MOST of your acquaintances have?" (See KENWERK in Appendix A).

3.2 Preliminary analysis

Essentially I will run a number of straightforward analyses for both saving measures. As both measures are ordinal, my main vehicle will be ordered probit. Since the explanatory variables are usually either categorical or ordinal, they will usually be recoded as dummies. In this section I treat all explanatory variables as if they were measured on an interval scale. The left hand side variable is treated either as an ordinal variable (in which case I use ordered probit) or as a variable measured on an interval scale. Table 12 shows some preliminary estimates based on either regression or ordered probit. Both the signs and the order of magnitude of the coefficients are the same across saving measures and across estimation methods. Household income appears to have a strong positive effect on saving, whereas a bigger family size tends to decrease saving. Highly educated people save more than people with less education. If income went up over the past twelve months, one is more likely to have saved than in the case where income fell. Assuming that not all income changes are fully anticipated, this is what a life cycle permanent income hypothesis would predict. For the purpose of this paper, the most striking outcome is that having friends or acquaintances with high incomes tends to reduce saving.

4. Empirical results

The results presented in Table 12 may be suggestive, but the underlying analysis is flawed in several ways: the right hand side variables are generally not measured on an interval scale; there is the possibility of individual effects being correlated with explanatory variables included in the equation; since the same respondent may appear in the analysis more than once, errors are most likely not iid, etc. The empirical strategy used in this section aims at remedying most of these problems. I will stick to the practice of running both ordered probits and regressions. The regressions are of the random effects type, while allowing for the possibility of correlation of individual effects with explanatory variables. To be more precise: the regression models estimated are of the form:

$$y_{it} = x_{it}'\beta + u_i + \epsilon_{it} \qquad (13)$$

where ϵ_{it} is assumed to be an iid error term and u_i an individual effect. The individual effect may be correlated with the explanatory variables on the right hand side of (13). We model this correlation by writing u_i as a linear function of the individual group mean of the explanatory variables plus an iid error term:

$$u_i = \overline{x}_{i\cdot}\,\gamma + \nu_i \qquad (14)$$

It is well-known that if we leave γ in (14) unrestricted, estimation of (13) with the fixed effects estimator yields exactly the same estimates for β as applying a random effects estimator to:

$$y_{it} = x_{it}'\beta + \overline{x}_{i\cdot}\,\gamma + \epsilon_{it} + \nu_i \qquad (15)$$

,cf. Mundlak (1978). Not imposing any restrictions on the parameters γ may induce considerable efficiency loss. My strategy will therefore be to restrict (14) (or equivalently, (15)), but apply Hausman tests (Hausman (1978)) throughout to verify the hypothesis that the ν_i in (15) are uncorrelated with the explanatory variables. The form of the restrictions is straightforward: I include in the vector of individual means all variables referring to the individual him- or herself, but not variables referring to his/her reference group.

In other words, it is assumed that unobserved individual heterogeneity in the propensity to save may be related to individual characteristics, like education, age, etc. but not to reference group variables, like the age and education of one's friends. As mentioned, the assumption can be tested in a straightforward way.

The modelling strategy for the ordered probits is essentially the same, except for the fact that there is not a straightforward Hausman test available to test orthogonality. In that case our test of orthogonality of the individual effects and the included explanatory variables amounts to inclusion of the individual mean of reference group variables (in addition to the means of the individual variables) and testing whether their coefficients are zero.

Table 13 shows results for a specification similar to the one presented Table 12. However, I now also control for net household wealth (the sum of all assets minus the sum of all liabilities: altogether some 50 types of assets and liabilities are distinquished). Household wealth has been transformed by the hyperbolic sine:

$$h(z_{it}) = \ln\left(z_{it} + \sqrt{z_{it}^2 + 1}\right) \qquad \textbf{(16)}$$

Furthermore, the means of all individual characteristics are now included. For the ordered probits the standard errors are corrected for the fact that the same individuals appear in the data more than once.

Perhaps the most striking aspect of including the individual means is the effect of net household wealth. Household wealth itself does not have an effect on saving, but it correlates strongly with the individual effect. An obvious interpretation is that people who tend to save a lot (other things being equal) also have high wealth (and not the other way around). A similar comment can be made with respect to household income: people who save a lot also have higher incomes. But here in addition to the correlation with the individual effect, household income has an independent effect on saving, as would be predicted by a standard model of consumption smoothing. Interestingly, whether or not income rose over the past 12 months now looses its significance. Similarly, education has lost its significance. Also household size, which appeared to be quite significant in Table 12, is completely insignificant: people who save a lot are also the people with small households. But once this is taken into account, there is no independent household size effect left. Age and birth year are not significant, although here the implausible linear specification may be to blame (see below). Occupation in one's reference group appears to be an important determinant of an individual's propensity to save.

The variable of main interest in this paper, reference group income remains significant in all specifications.

Table 14 shows the results of including the same variables as in Table 13, but now with all ordinal or categorical variables in dummy form. Moreover, age is entered as a fifth degree polynomial and birth year as a cubic. I have not included time effects for several reasons: first of all, there is an identification problem: one cannot separately identify age, time and birth year effects. Secondly, there is no obvious reason why savings should exhibit a time effect (in contrast to wealth). In any case, age and cohort effects are of no central concern

in this analysis. To avoid an ocean of hard to interpret numbers I only report estimates if a group of dummies is significant at at least the 5% level. Thus the variables for which no parameter estimates are reported, are included in the analysis, but the corresponding estimates are not shown. The age of people in one's reference group is never significant and hence does not appear in the table at all. The same is true of the respondent's own age.

Own household income and reference group income is always significant. Own household income dummies are always significant at levels of less than 0.0000; the reference group income dummies are significant at levels of 0.002 or below. The estimated dummy coefficients are displayed in Figures 1 and 2, along with 95% confidence intervals (the vertical bars). One notices that the effects of the dummies for own household income increases monotonically with income. The effect of the reference group income dummies *falls* almost monotonically with the level of the income class they refer to. The ordered probit analyses produce more accurate estimates than the random effects regression, and hence more significant dummies. Age is never significant, and birth year only in one specification.

5. Conclusions

Although the theoretical exercise in Section 2 did not generate unambiguous predictions as to the exact effect of reference group income on savings, it did point at a number of possible parameter configurations that would generate a depressing effect of reference group income on savings. The specifications estimated in the previous section suggest that such a depressing effect may indeed exist.

The weakest point of the analysis so far probably is the possible endogeneity of reference group incomes. This may come about in at least three different ways:

1. People who spend more tend to associate with others who have higher incomes.
2. Big spenders get together. By spending more, higer incomes are signaled and what these respondents in the survey report is really higher spending in their reference group rather than higher incomes.
3. People who spend more justify this (perhaps subconsciously) by stating that their friends and acquaintances have higher incomes.

Appendix A: Text of the most relevant questions

Below follows the text of a number of questions of which the answers have been used in the empirical analysis. The names of the corresponding variables are given in capitals at the beginning of each question:

INKHH

The TOTAL NET INCOME OF YOUR HOUSEHOLD consists of the income of all members of the household, after deduction of taxes and premiums for social insurance policies, taken as the sum total over the past 12 months. Into which of the categories mentioned below did the total net income of your household go IN THE PAST 12 MONTHS? If you really don't know, type 0 (zero).

0 don't know
1 less than Dfl. 20,000
2 20,000 - 28,000
3 28,000 - 43,000
4 43,000 - 80,000
5 80,000 - 150,000
6 150,000 or more

INKNORM

Is this income (the net income of your household that you have just mentioned) unusually high or low compared to the income you would expect in a 'regular' year, or is it regular?

1 unusually low
2 regular
3 unusually high
4 don't know

INKROND

How well can you manage on the total income of your household (as mentioned in the second from last question)?

1 it is very hard
2 it is hard
3 it is neither hard nor easy
4 it is easy
5 it is very easy

INKEVEN

Over the past 12 months, would you say the expenditures of your household were higher than the income of the household, about equal to the income of the household, or lower than the income of the household?

1 the expenditures were higher than the income
2 the expenditures were about equal to the income
3 the expenditures were lower than the income

INKAANK

Was the purchase of a house or car, or were other (big) investments part of these expenditures?

1 yes
2 no

INKSEC

When you ignore the purchase of a house or car, or other (big) investments, would you say the expenditures of your household, over the past 12 months, were higher than the income of the household, about equal to the income of the household, or lower than the income of the household?
1 the expenditures were higher than the income
2 the expenditures were about equal to the income
3 the expenditures were lower than the income

INKVER

The TOTAL NET INCOME OF YOUR HOUSEHOLD consists of the income of all members of the household, after deduction of taxes and premiums for social insurance policies, taken as the sum total over the past 12 months. Compared to about one year ago, did the total net income of your household increase, remain about the same, or decrease?
1 increased
2 remained about het same
3 decreased

INKTOE

The TOTAL NET INCOME OF YOUR HOUSEHOLD consists of the income of all members of the household, after deduction of taxes and premiums for social insurance policies, taken as the sum total over the past 12 months. Do you think, taking into account possible changes within the household, the total net income of your household will increase, remain the same, or decrease, IN THE NEXT 12 MONTHS?
1 increaseINKTOEHO
2 remain the sameINKZEKER
3 decreaseINKTOELA

OPZIJ

Did you put any money aside IN THE PAST 12 MONTHS?
1 yes
2 no

HOEVOPZY

About how much money has your household put aside IN THE PAST 12 MONTHS? If you really don't know, type 0 (zero).
1 less than Dfl. 3,000
2 3,000 - 10,000
3 10,000 - 25,000
4 25,000 - 40,000
5 40,000 - 75,000
6 75,000 - 150,000
7 150,000 or more

ERFENIS

Do you expect to be left a substantial inheritance in the future, or do you expect to receive a substantial gift in the future?
1 yes
2 no

0 don't know

ERFGELD

Will these inheritances and/or gifts in total be worth a great amount of money, an average amount, or a small amount?

1 great amount
2 average amount
3 small amount

ERFDENK

Have you ever thought about leaving a bequest?

1 yes
2 no

ERFBEL

Some people think it important to leave a bequest to their children or to other heirs, while other people don't find that important. Do you think this is important, or not?

1 very important
2 important
3 neither important nor unimportant
4 unimportant
5 very unimportant

TESTA Have you made a will?

1 yes
2 no

GELUKKIG

All in all, to what extent do you consider yourself a happy person?

1 very happy
2 happy
3 neither happy nor unhappy
4 unhappy
5 very unhappy
6 don't know

The following questions concern your circle of acquaintances, that is, the people with whom you associate frequently, such as friends, neighbors, acquaintances, or maybe people at work.

KENLTD

If you think of your circle of acquaintances, into which age category do MOST of these people go? Please select the answer that is closest to reality.

age (in years) is mostly:

1 under 16
2 16 - 20
3 21 - 25
4 26 - 30
5 31 - 35

6 36 - 40
7 41 - 45
8 46 - 50
9 51 - 55
10 56 - 60
11 61 - 65
12 66 - 70
13 71 or over

KENHH

The people in your circle of acquaintances may live alone or share a household with other people (for example with a partner and children). Of how many persons do MOST households of your acquaintances consist?

1 one person
2 two persons
3 three persons
4 four persons
5 five persons
6 six persons or more

KENINK

How much do you think is the AVERAGE total net income per year of those households?

1 less than Dfl. 17,500 per year
2 17,500 - 20,000
3 20,000 - 24,000
4 24,000 - 28,000
5 28,000 - 34,000
6 34,000 - 43,000
7 43,000 - 60,000
8 60,000 - 80,000
9 80,000 - 105,000
10 105,000 - 150,000
11 150,000 or more
0 don't know

KENOPL

Which level of education do MOST of your acquaintances have?

1 primary eduaction
2 junior vocational training
3 lower secondary education
4 secondary education/pre-university education
5 senior vocational training
6 vocational colleges/first year university education
7 university education

KENWERK

What kind of employment do MOST of your acquaintances have?

1 self-employed

2 practicing a free profession
3 working in the family business
4 employed on a contractual basis
5 mostly no paid job

MANUUR

If you think of the MEN among your acquaintances, how many hours per week do they work on average?
number of hours:

VROUWUUR

If you think of the WOMEN among your acquaintances, how many hours per week do they work on average?
number of hours:

References

Binder, M. and H. Pesaran (forthcoming), "Life-Cycle Consumption Under Social Interactions", Journal of Economic Dynamics and Control, forthcoming.

Brock, W.A. and S.N. Durlauf (forthcoming), "Interactions-Based Models", Handbook of Econometrics (J.J. Heckman and E. Leamer, eds.).

Kapteyn, A., S.A. van de Geer, H. van de Stadt, and Tom Wansbeek (1997), "Interdependent Preferences: an Econometric Analysis", Journal of Applied Econometrics, 12, 665-686.

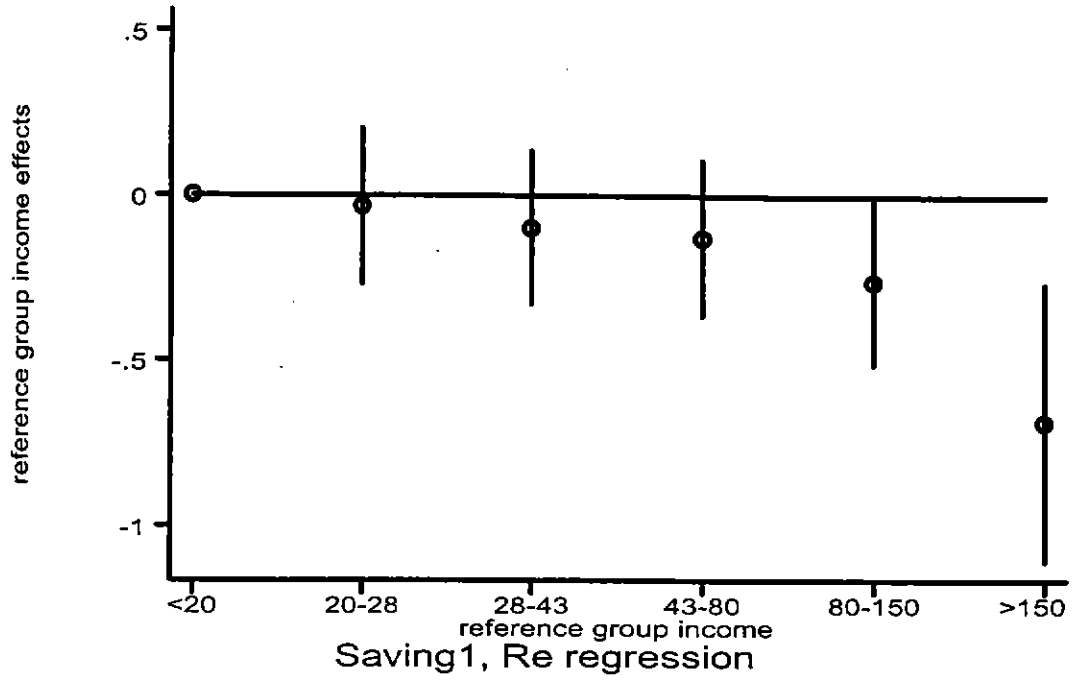
Manski, C. (1993), "Identification of Endogenous Social Effects: The reflection Problem", The Review of Economic Studies, 60, 531-542.

Mundlak, Y. (1978),"On the Pooling of Time Series and Cross Section Data," Econometrica, 46, 69-85.

Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting", Quarterly Journal of Economics.

Van de Stadt, H., A. Kapteyn, and S.A. van de Geer (1985), "The Relativity of Utility: Evidence from panel data", The Review of Economics and Statistics, 67, 179-87.

Woittiez, I. and A. Kapteyn (1999), "Social Interactions and Habit Formation in a Model of Female Labour Supply," Journal of Public Economics.



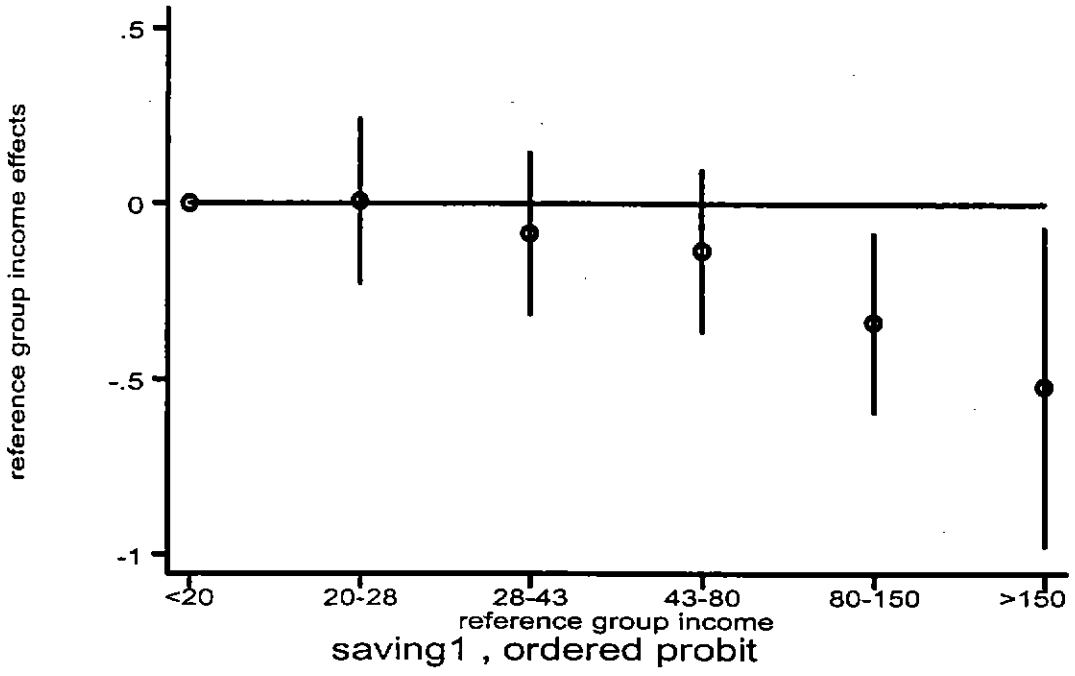
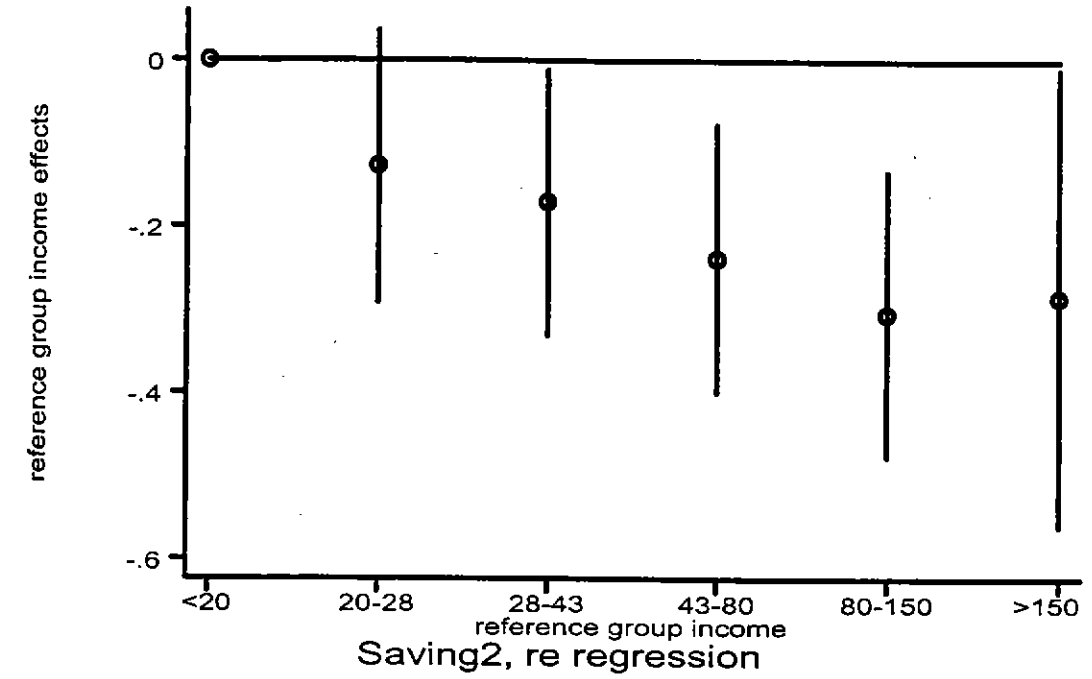


Figure 1: Estimates of reference group income dummy effects on saving1



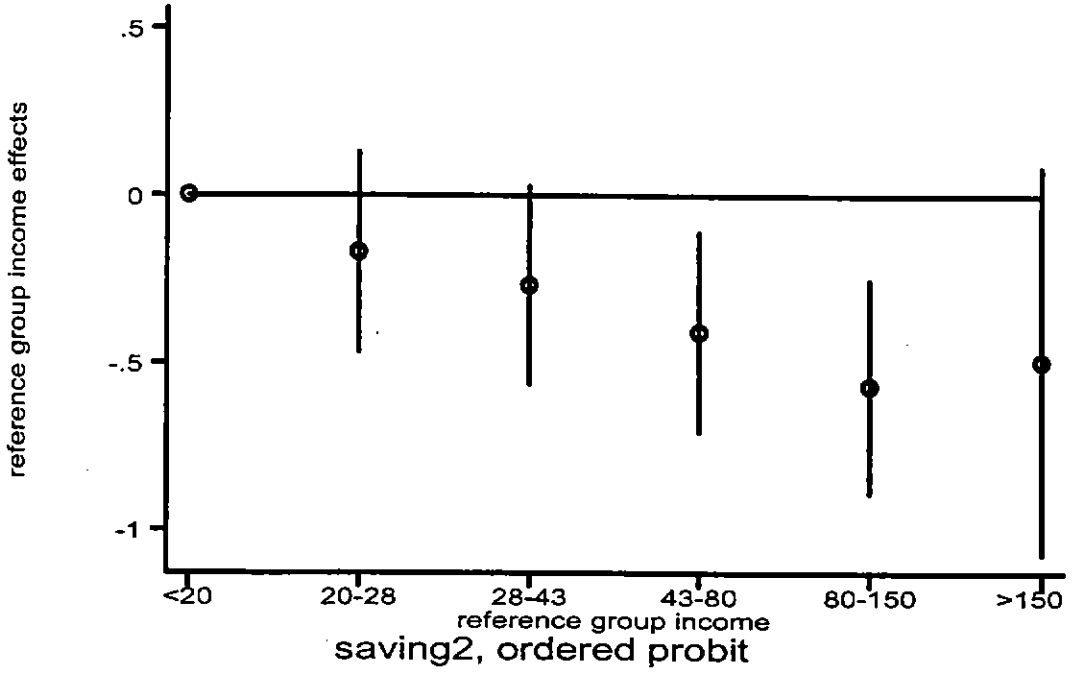


Figure 2: Estimates of reference group income dummy effects on saving2

Table 1: Frequency distribution of reported average reference group incomes

Average annual income acquaintances	Freq.	Percent	Cum. Percent
Less than Dfl.17,50	353	4.1	4.1
17,500-20,000	116	1.4	5.5
20,000-24,000	189	2.2	7.7
24,000-28,000	368	4.3	12.00
28,000-34,000	730	8.5	20.5
34,000-43,000	1371	16.0	36.5
43,000-60,000	2554	29.8	66.4
60,000-80,000	1790	20.9	87.3
80,000-105,000	779	9.1	96.4
105,000-150,000	249	2.9	99.3
150,000 or more	58	0.7	100.0
Total	8557	100.0	

Note: Amounts refer to annual incomes; Dfl. 1 is approximately US$ .50.

Table 2: Frequency distribution of self-reported after tax household incomes

Estimated total net income of hh	Frequency	Percent	Cum. Percent
Less than Dfl. 20,000	679	4.1	4.1
20,000-28,000	1358	8.2	12.2
28,000-43,000	3473	20.9	33.1
43,000-80,000	7213	43.3	76.4
80,000-150,000	3623	21.8	98.2
150,000 or more	304	1.8	100.0
Total	16650	100.0	

Note: Amounts refer to annual incomes; Dfl. 1 is approximately US$ .50

Table 3: Cross tabulation of reference group incomes with own household incomes
(percentages are relative to row totals, except in the last column where they are relative to the column total)

Total net income of hh	*Total net income of households in reference group* (Dfl. 1000)						
	<20	20-28	28-43	43-80	80-150	150-	Total
< Dfl.20,000	89	42	48	41	7	3	230
%	38.7	18.3	20.9	17.8	3.0	1.3	2.9
20,000-28,000	53	191	150	58	7	1	460
%	11.5	41.5	32.6	12.6	1.5	0.2	6.0
28,000-43,000	39	144	859	347	30	2	1421
%	2.7	10.1	60.5	24.4	2.1	0.1	18.4
43,000-80,000	30	85	771	2455	178	7	3526
%	0.8	2.41	21.9	69.6	5.0	0.2	45.7
80,000-150,000	8	9	105	1148	646	14	1930
%	0.4	0.5	5.4	59.5	33.5	.7	25.0
150,000 or more	1	2	7	36	80	17	143
%	0.7	1.4	4.9	25.2	55.9	11.9	1.9
Total	220	473	1940	4085	948	44	7710
%	2.8	6.1	25.2	53.0	12.3	.6	100.0

Note: Amounts refer to annual incomes; Dfl. 1 is approximately US$ .50
$\chi^2(25)=5746$; rank correlation=.61

Table 4: Cross tabulation of reference group household size with own household size (percentages are relative to row totals, except in the last column where they are relative to the column total)

	Average family size in reference group						
Own family size	one	two	three	four	five	> 5	Total
one person	411	888	231	215	18	11	1774
%	23.2	50.1	13.0	12.1	1.0	0.6	11.4
two persons	224	3904	1017	872	43	20	6080
%	3.7	64.2	16.7	14.3	0.7	0.3	39.0
three persons	99	668	704	816	48	5	2340
%	4.2	28.6	30.1	34.9	2.0	0.2	15.0
four persons	158	350	601	2086	114	25	3334
%	4.7	10.5	18.0	62.6	3.4	0.8	21.4
five persons	84	137	222	929	202	24	1598
%	5.3	8.6	13.9	58.1	12.6	1.5	10.3
six or more	30	24	47	190	126	46	463
%	6.5	5.2	10.2	41.0	27.2	9.9	3.0
Total	1006	5971	2822	5108	551	131	15589
%	6.4	38.3	18.1	32.8	3.5	0.8	100.0

$\chi^2(25)=7368$; rank correlation=.52

Table 5: Frequency distribution of saving1

How much was put aside?	Frequency	Percent	Cum. Percent
No money put aside	5202	30.2	30.2
Less than Dfl. 3,000	3083	17.9	48.1
3,000-10,000	5445	31.6	79.7
10,000-25,000	2677	15.6	95.3
25,000-40,000	513	3.0	98.3
40,000-75,000	180	1.0	99.3
75,000-150,000	66	0.4	99.7
Dfl.150,000 or more	35	0.2	99.9
Total	17201	100	

Note: Amounts are in Dutch guilders (Dfl.); Dfl. 1 is approximately US$ .50.

Table 6: Cross tabulation of saving2 against whether people put money aside the past 12 months (row percentages in parentheses in the first two columns; column percentages in the last column)

Expenditures greater than income?	Did not put money aside	Put money aside	Total
Expenditures higher than income	646 (47.0)	729 (53.0)	1375 (13.2)
Expenditures about equal to income	1973 (40.5)	2899 (59.5)	4872 (46.6)
Expenditures lower than income	413 (9.8)	3787 (90.2)	4200 (40.2)
Total	3032 (29.0)	7415 (71.0)	10447

$\chi^2(2)=1277$

Table 7: Cross tabulation of saving1 against saving2 (row percentages in parentheses in the first three columns; column percentages in the last column)

Saving1: How much was put aside?	*Saving 2: Expenditures greater than income?*			
	Expend. higher than income	Expend. about equal to income	Expend. lower than income	Total
No money put aside	646 (21.3)	1973 (65.1)	413 (13.6)	3032 (30.4)
less than Dfl.3,000	283 (15.8)	1108 (61.7)	406 (22.6)	1797 (18.0)
3,000-10,000	307 (9.2)	1261 (37.9)	1762 (52.9)	3330 (33.4)
10,000-25,000	76 (5.4)	268 18.9)	1077 (75.8)	1421 (14.3)
25,000-40,000	14 (5.8)	41 (16.9)	187 (77.3)	242 (2.4)
40,000-75,000	6 (7.2)	18 (21.7)	59 (71.1)	83 (0.8)
75,000-150,000	1 (2.9)	13 (37.1)	21 (60.0)	35 (0.4)
Dfl.150,000 or more	3 (16.7)	1 (5.6)	14 (77.8)	18 (0.2)
Total	1336 (13.4)	4683 (47.0)	3939 (39.6)	9958

Note: Amounts are in Dutch guilders (Dfl.); Dfl. 1 is approximately US$ .50.
$\chi^2(14)=2319$; rank correlation=.45.

Table 8: Age distributions of respondents and of people in their reference groups

	Respondents' referencegroups			*Respondents*		
Age categories	Freq.	Percent	Cum. Perc.	Freq.	Percent	Cum. Perc.
under 16	201	1.3	1.3			
16-20	827	5.3	6.6	967	6.2	6.2
21-25	555	3.6	10.2	425	2.7	8.9
26-30	1031	6.6	16.8	928	6.0	14.9
31-35	2018	12.9	29.7	1457	9.4	24.2
36-40	2427	15.6	45.3	1766	11.3	35.6
41-45	2165	13.9	59.2	1972	12.6	48.2
46-50	2049	13.1	72.3	1874	12.0	60.2
51-55	1343	8.6	80.9	1498	9.6	69.8
56-60	1181	7.6	88.5	1233	7.9	77.7
61-65	956	6.1	94.6	1125	7.2	85.0
66-70	551	3.5	98.2	942	6.0	91.0
71 or over	286	1.8	100.0	1403	9.0	100.0
Total	15590	100.0		15590	100.0	

Table 9: Cross tabulation of respondents' education with education in their reference group (numbers on each second row are percentages of row total, except in the last column where they are percentages of the column total).

Education of respondents	*Education in reference group*							
	Prim.	Jun. voc.	Lowersec.	Sec.	Sen. voc.	Voc. coll.	Univ.	Total
Primary	81	162	198	331	80	44	20	916
	8.8	17.7	21.6	36.1	8.73	4.8	2.2	6.7
Junior vocational	63	534	521	493	251	53	7	1922
	3.3	27.8	27.1	25.6	13.1	2.76	0.4	14.0
Lower secondary	31	107	667	817	363	214	34	2233
	1.4	4.8	29.9	36.6	16.3	9.6	1.5	16.3
Secondary	15	28	187	814	250	486	189	1969
	0.8	1.4	9.5	41.3	12.7	24.7	9.6	14.4
Senior vocational	33	112	329	672	624	267	19	2056
	1.6	5.4	16.0	32.7	30.4	13.0	0.9	15.1
Vocational colleges	26	24	123	541	487	1725	218	3144
	0.8	0.8	3.9	17.2	15.5	54.9	6.9	22.9
University	12	3	17	151	78	534	665	1460
	0.8	0.2	1.2	10.3	5.3	36.6	45.6	10.7
Total	26	970	2042	3819	2133	3323	1152	13700
	1.9	7.1	14.9	27.9	15.6	24.3	8.4	

$\chi^2(36)=8963$; rank correlation=.58

Table 10: Did income rise, stay the same, or did it fall over the past 12 months? Frequency distribution of answers

Did income rise?	Freq.	Percent	Cum. Percent
Income decreased	1946	10.8	10.8
Income stayed the same	12026	66.9	77.7
Income increased	4004	22.3	100.0
Total	17976	100.0	

Table 11: Modal occupation in respondents' reference groups

Employment acquaintances	Freq.	Percent	Cum. Percent
Self-employed	1170	7.5	7.5
Professional	648	4.2	11.7
Working in the family business	142	0.9	12.6
Employee	11350	72.8	85.4
Mostly no paid job	2280	14.6	100.0
Total	15590	100.0	

Table 12: Regressions and ordered probits for two savings measures								
Dependent variable:	Saving1				Saving2			
Type of analysis:	Regression		Ordered Prob.		Regression		Ordered Prob.	
Explanatory variables	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Income in reference group (KENINK)	-.03	-2.8	-.03	-3.1	-.03	-4.5	-.06	-4.4
Household size (AANTALHH)	-.10	-8.4	-.09	-8.5	-.08	-9.9	-.13	-10.0
Education (7 levels; 1 is the lowest)	.05	5.5	.05	5.4	.01	1.4	.02	1.5
Household income (INKHH)	.52	27.6	.51	27.1	.21	16.7	.36	16.2
Did income rise? (2-INKVER)	.09	3.4	.09	3.6	.08	4.9	.15	4.9
Age	-.02	-1.5	-.02	-1.7	-.02	-2.6	-.04	-2.5
Year of birth	-.01	-0.47	-.01	-0.61	-.02	-2.5	-.04	-2.4
Modal age in ref. group (KENLTD)	.004	0.31	.005	0.48	.01	1.1	.01	0.94
Modal occ. in ref. gr. (KENWERK)	.04	2.9	.04	2.9	.02	2.1	.03	2.0
Constant	10.9	0.48			45.4	2.5		
R^2 or pseudo-R^2	0.18		0.06		0.09		0.05	
Number of observations	6706		6706		5047		5047	

Table 13: Random effects regressions and ordered probits for two saving measures
(with individual means of personal characteristics included to account for correlated individual effects)

Dependent variable:	Saving1				Saving2			
Type of analysis	RE regression		Ordered Prob.		RE regression		Ordered Prob.	
Explanatory variables	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Income in reference group	-.03	-2.9	-.04	-3.7	-.03	-3.8	-.06	-4.3
Household size	.02	0.34	.016	0.28	.03	0.7	.05	0.72
Individual mean hh size	-.14	-2.5	-.13	-2.3	-.12	-2.5	-.20	-2.9
Education	-.12	-1.1	-.10	-0.99	-.12	-1.9	-.21	-2.7
Individual mean education	.17	1.6	.15	1.4	.13	2.1	.23	2.9
Household income	.18	5.3	.18	5.0	.15	5.4	.26	4.9
Individual mean hh income	.31	7.6	.33	7.9	.02	0.69	.05	0.90
Incrise: Did income rise?	.06	1.7	.06	1.6	.04	1.6	.07	1.6
Individual mean of incrise	.03	0.57	.05	0.95	.05	1.4	.11	1.7
Age	.0007	0.05	.002	0.19	-.01	-0.70	-.01	-0.50
Individual mean of age	-.02	-0.74	-.03	-1.3	-.03	-1.6	-.05	-1.6
Birth year	.0003	0.02	-.009	-0.53	-.03	-2.5	-.06	-2.4
Modal age in ref. group	.002	0.19	-.0004	-0.03	.004	0.55	.005	0.34
Modal occupation in ref. gr.	.04	2.4	.05	2.9	.03	2.7	.05	2.5
Household net worth	-.002	-0.46	-.002	-0.36	.004	0.95	.006	1.0
Individ. mean of hh net worth	.04	6.9	.05	6.8	.02	4.0	.03	4.3
Hausman test for random eff.	$\chi^2(10)$= 12.2, p=.27				$\chi^2(9)$= 4.12 p=.90			
Number of observations	6573		6573		4946		4946	

Table 14: Random effects regressions and ordered probits for two saving measures; dummies used wherever appropriate
(with individual means of personal characteristics included to account for correlated individual effects; empty cells mean that the corresponding group of variables was not significant at the 5% level. The variables have been included in the analysis, but their parameters are not reported, to facilitate reading of the table.)

Dependent variable:	Saving1				Saving2			
Type of analysis	RE regression		Ordered Prob.		RE regression		Ordered Prob.	
Explanatory variables	Coef.	t	Coef.	t	Coef.	t	Coef.	t
hh income:20,000-28,000	.24	1.3	.29	1.3	.54	3.9	.92	4.1
28,000-43,000	.48	2.6	.54	2.4	.75	5.4	1.30	5.3
43,000-80,000	.53	2.8	.59	2.6	.85	6.1	1.48	5.9
80,000-150,000	.72	3.7	.80	3.4	.92	6.2	1.61	6.0
>150,000	1.14	4.8	1.12	4.00	1.0	5.2	1.80	4.3
Mean hh inc: 20-28,000	-.20	-0.88	-.15	-0.58			-.62	-2.5
28,000-43,000	.005	0.02	.06	0.23			-.57	-2.2
43,000-80,000	.37	1.6	.46	1.8			-.39	-1.4
80,000-150,000	.73	3.2	.83	3.1			-.22	-0.77
>150,000	1.11	3.7	1.30	3.5			-.31	-0.64
Income in ref. gr. 20-28	-.03	-0.27	.01	0.07	-.13	-1.5	-.17	-1.1
28,000-43,000	-.10	-0.83	-.08	-0.72	-.17	-2.1	-.27	-1.8
43,000-80,000	-.13	-1.1	-.13	-1.1	-.24	-2.9	-.41	-2.7
80,000-150,000	-.26	-2.1	-.34	-2.6	-.30	-3.5	-.57	-3.5
>150,000	-.69	-3.2	-.52	-2.2	-.28	-2.0	-.49	-1.7
Education: jun. vocational			-1.18	-2.6			.61	1.4
Lower secondary			-.38	-0.68			.72	1.6
Secondary			.043	0.01			.82	1.2
Senior vocational			-.65	-1.1			-.08	-0.17
Vocational college			-1.10	-1.5			1.1	1.8
University			-.71	-0.85			-.26	-0.41
Mean of dum. for jun. voc.			1.11	2.3	-.40	-0.98	-.71	-1.6

Lower secondary			.43	0.77	-.47	-1.1	-.88	-1.8
Secondary			-.15	-0.17	-.50	-0.69	-1.11	-1.6
Senior vocational			.66	1.1	.02	0.04	.028	0.05
Vocational college			1.19	1.6	-.58	-0.96	-1.20	-1.94
University			.87	1.0			.21	0.31
Mean of d. for hh size=2							-.36	-1.4
Household size=3							-.67	-2.4
Household size=4							-.81	-2.5
Household size=5							-1.13	-3.1
Household size>6							-.75	-1.9
Occ. in r.g.: professional	-.054	-0.66	.06	0.55				
Working in fam. business	-.10	-0.61	-.01	-0.07				
Employee	.12	2.2	.21	2.8				
Mostly no paid job	.05	0.70	.14	1.5				
Inc.. fell over the p. 12 m.	.03	0.56	.04	0.63				
Inc. rose over the p. 12 m.	.14	3.0	.14	3.1				
Birth year								
Birth year squared					.00	1.3		
Birth year cubed					-6.7e-09	-1.7	-5.29e-09	-2.5
Hyp. sine of hh wealth	-.002	-0.45	-.003	-0.39	.005	1.2	.008	1.2
Mean hyp. sine hh wealth	.045	7.0	.052	7.0	.017	3.5	.029	3.6
Hausman test for random eff.	$\chi^2(52)= 42.5$, p=.82				$\chi^2(52)= 31.8$ p=.99			
Number of observations	6573		6573		4946		4946	