THE VSB-CENTER SAVINGS PROJECT: DATA COLLECTION METHODS, QUESTIONNAIRES AND SAMPLING PROCEDURES

By Ellen Katrine Nyhus

Tilburg University
THE VSB-CENTER SAVINGS PROJECT: DATA COLLECTION METHODS, QUESTIONNAIRES AND SAMPLING PROCEDURES

By Ellen Katrine Nyhus

November 1996

The VSB savings project progress reports facilitate a quick dissemination of results obtained by researchers involved in the project. Essentially two types of papers may appear in the series. The first type reports on data analysis, describes cleaning procedures, sampling frames, etc. These help to give researchers in the project a good background to start using the data for their own analyses. The second type of papers presents results of academic research into savings. Often, but not necessarily, these papers will use the data made available in the savings project for fundamental research. Frequently these papers also appear in the CentER DP series.

Apart from the papers appearing in the CentER discussion paper series, progress reports are not refereed. Readers of the progress reports are encouraged to communicate comments and criticisms to the authors of the reports.
Preface

I have written this report hoping it will be useful for other researchers using the data collected for the VSB-CentER Savings Project. The report has been written during my stays as a visiting researcher at CentER, and I wish to thank the staff and students at CentER for their support and hospitality. They make CentER an enjoyable and stimulating place to work. I also would like to thank CentER for financial support, making the visits possible.

In particular, I want to thank Herman Cunphuis (CentER), Adriaan Hoogendoorn (University of Amsterdam), Maarten Ketelaars (CentER) for providing me with information about the sampling procedures and the panels. Without their help, this report could not have been written. I also would like to thank Kjell Grunhaug, Jonas Gunnarson, Stefan Hochgürtel, Arie Kapteyn, Arthur B. Kennickell, and Karl-Erik Wärneryd for helpful comments on earlier drafts of this report. Georgina Broeken deserves a thank you for polishing my English. All remaining mistakes and inadequacies are my own responsibility.

This report does not contain a complete discussion of sample biases and possible effects of the data collection method. There are still many issues that should be evaluated and discussed concerning the data and the conclusions we can draw from analyses of them. If anyone has comments to the points already discussed here or suggestions about other issues that should be included in this report, I will be glad to hear from you. My address is:

Ellen Katrine Nyhus
Institute of marketing/NHH
Breiviken 2
N-5035 Bergen- Sandviken
Norway
e-mail: ellen.nyhus@nhh.no
I. Introduction

The VSB-CentER Savings Project was started in 1990. The goals of the project are 1) to test the descriptive and predictive power of economic and psychological factors on households' saving behaviour and 2) to study the stability of any possible influence of these variables on saving behaviour. The project is expected to last 5 - 7 years.

Data will be collected every year from a panel consisting of about 3000 households. Data are collected by computer assisted panel research. The first and second round of data collection took place in 1994, the third was carried out in 1995, and the fourth in 1996. Data about financial matters, work situation, demographics, psychological traits, health, and other living conditions are collected. The data set is unique, as it allows studies of both psychological and economic factors' effects on economic behaviour. In most large research projects on economic behaviour, psychological variables are not measured and their influence on economic behaviour cannot be tested.

No data analysis should be carried out before an assessment of the data accuracy and validity has taken place. The purpose of this report is to describe and evaluate the data collection methods and sampling procedures applied by the fieldwork company, "Stichting TelePanel" (the TelePanel Foundation). Strengths and weaknesses related to the methods are discussed with respect to external and internal validity. External validity refers to which degree results from one study can be applied to other populations beyond the sample, while internal validity refers to which degree one can be sure that the observed variance in a dependent variable is caused by the independent variable (Cook & Campbell, 1979). Many factors might represent threats to internal and external validity, and in this paper I focus on possibilities of sampling and nonsampling errors. These possible threats to validity should be considered before drawing conclusions and generalising results from analyses based on the Savings Project data.

The second section of this report focuses on possibilities of nonsampling errors. It contains a description of the applied data collection method as well as a discussion of advantages and disadvantages associated with it. The different parts of the questionnaire are briefly described and key informants within each household are identified. Results from previous validation studies are reported, as well as some comments from the respondents of the 1st wave of data collection. These reports are used as a basis for discussing the accuracy we might expect from the data collected for the Savings Project. Section three focuses on possibilities of sampling errors. It contains a description and evaluation of the sampling procedures applied for recruiting panel members. The report ends with a summary in section four. The report is based on information provided by staff at CentER and at Stichting TelePanel in Amsterdam.
2. Data Collection Procedures and Data Quality

In this section the data collection procedures and the questionnaire used for the VSB-CentER Savings Project are in focus. Advantages and disadvantages related to the data collection procedure will be discussed. The different parts of the questionnaire are shortly described, as well as the selection of key informants within each household. The section also contains a note on how respondents have answered questions about financial matters in previous investigations. As no validation checks have been performed with respect to the responses to the Savings Project questionnaires, experience from previous investigations is used as a base for the discussion of the data quality we might expect from the data collected for the project. Some comments from the respondents participating in the 1st wave of data collection are also included.

2.1 Computer Assisted Panel Research (CAPAR)
The data are collected by Stichting TelePanel, University of Amsterdam. A telepanel is a panel of households connected to a central computer via modems. Questions are sent out via the modems at the request of researchers and organisations, and the respondents answer them in their homes during computer sessions. The data programmes used for administering the questionnaires and responses are user friendly, so that it should be possible for all kinds of respondents to complete the questionnaires without assistance. Participants have agreed in telephone recruitment interviews to become members of the panel and regularly complete questionnaires. Households that are selected for participation get a PC and modem installed in their homes. They can use the PC for private purposes as long as they fill out the questionnaires they receive. Some panel members have their own PC and in these cases they use their own PC instead of the PCs belonging to the Stichting TelePanel. I do not know whether these panel members receive some other type of reward in return for their participation in the panel.

The sample used for the VSB-CentER Savings Project consists of two sub-samples. One sub-sample is intended to be representative of the Dutch population with respect to certain socio-economic variables. This panel (the Stichting telepanel - or the STP) consists of individuals in about 2000 households. The other sub-sample is a sample of Dutch high-income households, consisting of individuals in about 1000 households (the High Income Panel - or the HIP).

The STP members receive questions every Fridays, and they are supposed to complete them during computer sessions during the following weekend. The intention is that no session will last more than 30 minutes. The responses are transmitted to Stichting TelePanel's central computer. Panel members who do not answer within 3 days (the following Monday), are called and asked to fill in the questionnaire before the following Wednesday. If the household members still do not respond, they are counted as nonresponse. Respondents who have problems with any of the questions can call a help desk that has weekend service. In addition, there are technical staff available if there are problems related to the computers or the telephone lines.

The HIP members receive questionnaires once a month. They get longer questionnaires than the STP members, but the total number of questions is lower. No session is supposed to last more than an hour. The HIP members also get more time to answer the questionnaires than the STP members.

Computer assisted interviewing is different from the traditional interviewing techniques in many ways. The questionnaires are reproduced by a computer, following instructions from the researcher and a programmer. The interviewer or respondent enters the answers and the data can be processed immediately with a minimum of interface (Sikkel, 1988). Members of a telepanel answer questions on a regular basis and supervise the interviewing themselves by starting the data programme and answering the questions that are presented to them by typing their responses. This data collection method is called Computer Assisted Panel Research (CAPAR). As all data collection methods, CAPAR has advantages and disadvantages. These will be described below.

Advantages
The main advantages of the CAPAR are that it is both cheaper and faster than other data collection methods, and the data collectors save on paperwork as the coding of the data can be programmed and carried out as soon as responses are entered into the machine. Moreover, computerised interviewing allows complex routing and high control over question flows. Answers can be edited in the field by validation of responses, the questionnaires can be personalised and the order of the questions randomised. As there is no interviewer present, it avoids any interviewer effects and it is more convenient for the respondents as they can answer the questionnaires at times that suit them. Finally, the respondents might also improve in their ability to answer questionnaires over time, as they misunderstand less (Sikkel, 1988; Snijkers, 1992). These issues will be outlined in more detail below.

Costs
One advantage of the telepanel is that it is a relatively cheap way of collecting data. It is costly to establish a telepanel, as the computers that are distributed to the households cost a substantial sum of money. However, the data collection does not require any interviewers and the coding can be programmed and carried out by the computer. CAPAR is thus cheaper compared to paper and pencil interviews in the long run, since both the data collection and the data processing are cheaper (Sikkel, 1988).

Also the fact that data is collected from a panel might give cost advantages compared to cross-sectional data collection. The reason for this is that the differences in, for example, mean subsequent waves instead of two independent samples, Verwey et al (1989) argue that on the aggregate level the information about income obtained from a panel of 1000 households is comparable to that obtained from cross sectional research based on 8000 households. In general, this advantage of precision is higher, the lower nonresponse and attrition there is in the panel. Given that a researcher wants a certain level of accuracy, he/she needs fewer respondents...
**Time**

Collecting data through the telepanel has the advantage that it is an efficient interviewing method for researchers who are ready to collect data. A large number of respondents are available immediately and they have also expressed willingness to co-operate. Hence, time is saved as the recruitment of respondents has been carried out a priori. The teleinterviewing is also quicker than mail surveys, as delays due to the post system are avoided.

The coding of the data can be carried out as soon as the responses have been entered, and this makes the data available much sooner than with other data collection methods. When the answers are recorded on paper, data have to be punched in a separate operation which can be very time consuming.

Another advantage with teleinterviewing over telephone and personal interviews is that all interviews are carried out at the same point in time. When data are collected by personal or telephone interviews and the sample size is large, it might take several weeks to complete the interviews. Respondents will answer the questionnaires at different points in time, and this might cause differences in responses. For example, responses to questions about attitudes and expectations might be influenced by factors such as economic news, political discussions and whether there is a holiday coming up. These «history effects» are avoided with teleinterviewing.

**Possibilities of routing**

CAPAR has the advantage that complex routing structures can be built into the programme. Questionnaires can be tailored to various respondent groups. Compared to, for example, mail questionnaires, this has the advantage that respondents do not see the questions that he/she is not supposed to answer. For instance, some questions are meant for people answering "no" to a specific question, while other questions are relevant for the people that answered "yes". In mail questionnaires this can often be a problem because commands like "if you answered "yes" to the previous question, please, proceed to question 20" might be perceived as difficult and confuse the respondent. It might thus lead to mistakes or nonresponse. This problem is avoided by the computer aided interviewing as the routings are programmed. In addition, it reduces the workload of the respondent, which also might have a positive effect on data quality (Saris, 1991).

The possibilities for complex routing and programming also have the advantage that the routing condition can contain complicated calculations based on earlier answers. Answers to one question, like names and numbers, can be inserted in later questions. The order of the questions can also be randomised so it is possible to control for priming effects. However, if the routing structure is complex, this possibility is limited. The routing demands a certain order within blocks of questions.

**Possibilities for corrections**

In computerised interviewing it is possible to ask confirming questions, to make sure that the respondent answered correctly before the follow-up questions are presented. Range errors can be avoided by programming checks that can assess validity of answers, for example on dates (if people answer dates over 31 or month numbers higher than 12). Moreover, it is possible to program consistency checks directly by letting the programme compare responses to two different questions that are supposed to measure the same variable. The programme can repeat the questions or ask a third question on the same matter. Hence, computerised interviewing allows interactive editing which in most cases would not be possible in paper and pencil interviews (Saris, 1991). Data cleaning can to a certain extent be performed while the respondent is still available.

A mentioned earlier, panel research provides an extra source of information to check the validity of the data. As the same respondents will answer the same questions in subsequent waves, it is possible to clean data based on information from previous interviews (Saris, 1991). Missing data might in some cases be replaced with answers from earlier waves, and unlikely answers can be checked against previous answers to the same question. In the VSB-ComER Savings Project this way of improving data quality is used for data concerning financial data.

**Control for interviewer bias**

CAPAR is an impersonal form of interview. Question text is revealed on the screen and the respondents are given the opportunity to enter data. As the respondents do not have contact with any interviewers, biases due to effects of the interviewers are avoided. For example, it has been found that interviewers do not read questions exactly as they have been formulated. Moreover, interviewers do not always help the respondent in the same way if questions are misunderstood. A consequence of these differences might be that responses to questions collected by different interviewers are not comparable (Saris, 1991). When CAPAR is used, this problem no longer exist as the questions on the screen will be the same for all respondents.

**Convenience for respondents**

Another important advantage with CAPAR over telephone and personal interviews, is that the respondents can, to a certain extent, choose themselves when to answer the questionnaire. This can be particularly important when collecting financial data, because respondents will not feel the pressure from a waiting interviewer to give a quick answer. Especially in telephone interviews a long break in the interview while waiting for a question or an answer can be quite annoying for both interviewer and respondent (Snijkers, 1992). Many will need time to find the correct financial data, and teleinterviewing enables the respondents to check information by verifying their records or consulting with other members of the family. CAPAR therefore increases the probability of getting thoughtful replies instead of some hasty estimates given under pressure.

**Panel Conditioning**

The panel membership might have different effects on the respondents. One of these can be that the panel members gradually improve in answering questionnaires because over time, they obtain a better understanding of the surveys and thus restrict answers in ways consistent with this understanding. The Savings Project questionnaire can also make people more conscious about how much they save and earn so that the answers might be more accurate in later waves. If this is the case, the panel design might improve responses.
Disadvantages
Teleinterviewing also has some drawbacks. First of all, some of the saved costs and time in the data collection and data coding stages will be outweighed because the establishment of the panel is expensive and because the construction of the questionnaires takes more time with computerised interviewing. Moreover, the routing possibilities might cause serious mistakes in different ways if the programme does not operate correctly in every situation. As no person is present during the interview, there is little control over which person in the household actually answers the questionnaire. Finally, panel conditioning might have negative effect on the respondents' motivation for answering the questions accurately and the panel membership might even cause changes in the respondents' behaviour.

Time and costs
The teleinterviewing generates very high start-up costs in hardware, software and recruitment of panel members. This might be rather risky for the field workers. If the budget is overdrawn in the initial stages, the project might have to stop before any data are collected, and nothing is gained except a lot of computers and modems. The computers must also be maintained, and this might also cause high costs.

Designing and testing questionnaires for computerised interviewing is much more complicated than for paper and pencil interviews (cf. Saris, 1991). These stages are therefore both more time consuming and more expensive. A questionnaire designer has to write computer programmes which define every aspect of the interview in advance. As the routing makes the questionnaire quite inflexible when it is in use, it is important that all possible situations are taken care of in the programme so that the correct data is collected (Snijkers, 1992).

Routing
The routing in the computer aided interviewing has the drawback that a mistake made by the respondent when answering one question might cause a chain reaction of mistakes. For example, if a respondent gives a wrong answer to the question about whether he has one or more checking accounts, this might lead to a lot of missing data. For example, if a respondent by mistake answers that he does not have a checking account, he will not receive answers to the questionnaire. Finally, panel conditioning might have negative effect on the respondents' behaviour.

There is also a possibility of faulty routing if the routing structure becomes too complex. This can also lead to missing data, as some respondents do not get the questions they were supposed to get or they get questions that are not relevant for them. Too many checks for consistency may also cause that the interview programme runs very slowly and the interview can thus be interrupted too often by the programme (Snijkers, 1992).

Control with respondents
As with mail interviews, there is no control for who actually answers the questionnaire. It might happen that some of the household members do not have time to fill in the questionnaire and therefore leave it to some other household member in order to fulfill the obligation to Stichting TelePanel. There is also the chance that household members answer the questions together and not alone as intended. This might be an advantage for some of the questions on financial matters, but a disadvantage for questions on psychological variables.

Panel Conditioning
There is still little knowledge about how the respondents are affected by answering questionnaires from different researchers on a regular basis. There might for example exist priming effects, in the sense that a set of questions from one organisations might affect the different organisations collect data on the same matter in subsequent weeks, for example balances on bank accounts, this might cause respondents to become fed up with answering and simply skip these questions or the whole questionnaire.

Above we mentioned that people might get better in answering the questionnaires as they participate in the panel. However, the experience in answering questionnaires might also have negative effects. Panel members might become fed up with answering all the questionnaires and learn how to answer in order to minimize the work load (always choose the "I do not know - options" or become "careless respondents" (give random answers)). Inaccurate data might therefore be collected because panel members would like to keep the computer for a minimum of effort.

The panel membership itself may also cause that people change as well as their behaviour of interest to the researcher. Panel members might become more conscious of their participation in the panel as time goes by, which manifests itself in altered attitudes or behaviour (Ferber, 1966). Panel members will thus become atypical of the population group which they are meant to represent. For example, answering questions about expenditure and saving on a regular basis might make the respondents more conscious about their own spending behaviour and self controlling strategies. The questionnaire can also make them more aware of saving alternatives and reasons for saving, which might result in a change in behaviour. The repeated questions about financial data might even cause that answering as high saving as possible can become some sort of saving goal for some of the panel members. Every year they might want to report higher saving than the previous year.

2.2 Key Informants and the Questionnaire
Below a short description of the different parts of the questionnaire will be given as well as the key informants for these parts. The different parts of the questionnaire are called WRK (household, work and pension), INC (income and health), HS (accommodation and mortgages), WTH (assets and loans), and PSY (economic psychological variables) respectively. Some routing is reported here, but far from all. The routing system in the questionnaire is quite complicated and can be found in the questionnaires or in the variable documentation distributed from Coenen.

Selection of informants
Not all household members answered the questions in the savings project questionnaire. No household member below 16 years was asked to answer any questions. For some parts of the questionnaire, only one or two members of each household were asked to answer on behalf of the entire household. An overview of which persons in the households answered the different parts of the questionnaires, is given in table 2.1.

<table>
<thead>
<tr>
<th>Main Subject of Questionnaire</th>
<th>Key informants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. WRK - Household, Work, and Pension</td>
<td>All household members &gt; 15 years for whom the specific questions are relevant to (who have, had or are looking for a paid job). Some questions concerning children are only asked the heads of the households.</td>
</tr>
<tr>
<td>2. INC- Health and Income</td>
<td>Personal income: All household members &gt; 15 years for whom the specific questions are relevant.</td>
</tr>
<tr>
<td>3. HS - Accommodation and Mortgages</td>
<td>Household income: Heads of households and spouses/partners</td>
</tr>
<tr>
<td>4. WTH - Assets and loans</td>
<td>Household financial officer</td>
</tr>
<tr>
<td>5. PSY - Economic-psychological variables</td>
<td>Questions about personal properties to all panel members &gt; 15 years Questions about shared property should be answered by one household member only.</td>
</tr>
</tbody>
</table>

| Questions about income, income expectations, saving plans, saving motives, attitudes towards saving and debt, risk aversion, and changes in consumer expenditure were asked to the head of the households and spouses/partners only. (The questions about risk aversion were only asked to the household with net income above DH 20000 a year.) Questions about personality traits were asked all household members > 15 years. |

Table 2.1
Overview of Key Informants of the Households

Panel members were categorised according to their self reported position in the household, the categories being "head of household", "spouse", "unmarried partner", "parent" (in law), "child living at home", "house mate" and "other". The routing was based on these categories in addition to a category called the "household financial officer" (HFO), which is defined below. Some blocks of the questionnaires were routed only to the HFOs, others to the heads of the households, or to the heads of the households and their spouses/partners. Other sections were answered by all members of the household 16 years or older. Questions were only asked to those respondents they were relevant for, so questions about last paid job were not asked to those respondents that had their first paid job, and respondents who do not have a boat were not asked about the value of a boat. It is therefore difficult to say anything about how much time it would take each respondent to answer the questionnaire and also how many should have answered each question.

In the cases where only one household member is supposed to answer questions on behalf of the entire household, a strategy for selecting this household member must be chosen. The most important consideration in the selection of informants of household financial affairs is knowledge. Ideally, the household member with access to the information requested should answer the questions. For some questions, it is assumed that the heads of the households are the best informants. With respect to household consumption and assets, it is natural to expect that the person who usually does the shopping and pays the bills has the most knowledge about these matters. The person who is responsible for bills and other payments is for the purpose of this project called the "Household Financial Officer" (HFO). The procedures followed to identify the HFO, are outlined below.

The identification of the HFO took place some weeks before the VSB questionnaire was distributed (information about which household member is the HFO is not available in the data files). The panel members were first asked whether they were the person in the household that usually took care of bills and financial matters. If only one person within the household said that he or she usually took care of bills, this person was identified as the HFO. If more than one household member said that they took care of bills, selection among these was based on information about income. The person with the highest income was then identified as the HFO. If their income happened to be the same, selection was instead based on who of the persons had reported that he/she was the head of the household. In these cases, the head of the household was identified as the HFO.

The Questionnaire
The total questionnaire consists of five parts which were distributed to the respondents according to the agreements with the panel members. The STP receive questions in portions such that the questions could be answered in 30 minutes or less. The HIP receive the questionnaires in portions which would take an hour or less for them to complete. I do not know which parts of the questionnaires are sent out together, nor how many weekends the panel members answered the questionnaires of the Savings Project.

Each part of the questionnaire is described below. The order in which the five parts of the questionnaire was sent out is not known, so the order chosen for this presentation is not based on any such information. The description of the questionnaire is based on the questionnaire for the 1st wave of data collection. The questionnaires used for the 2nd, 3rd, and 4th wave contain additional questions.

Household, Work, and Pensions (WRK)
The first block of WRK, called "Block household", contains questions about primary occupation. All household members 16 years or older answer these questions. The responses to these questions decide what questions the respondents receive. The heads of the households are asked about marital status and children, and heads of households born before 1953 are also asked about grandchildren. Heads of households responding that they have children who no longer belong to the household, are asked about the date of birth and sex of these children, as well as the reasons for why they moved.
"Block work" is asked to respondents that currently have a paid job or that used to have a paid job. In the cases they have had more than one paid job, they are asked questions concerning the job on which they spent/spent most of their time. They are asked about what type of organisation they are/were employed in, how long they have had the paid job, their profession and position, how many hours they work/worked per week, and number of days off. They are then routed to the questions about pension.

The beginning of "Block Pension" contains questions to respondents that currently have a paid job or used to have a paid job. The former group is asked about whether they have any additional work to the paid job they already reported, how many hours per week they work at this additional work and how many hours they would like to work. The latter group is asked about when they stopped working, for what reasons and how much they earned from their last job. Thereafter, both groups are asked about which pension plans and retirement insurance they take/took part in.

All respondents are then asked whether they are looking for another job and if they are, what the reasons for this is. Respondents who are looking for another job are routed to "Block Looking for", which contains questions about this matter (why and for how long they are looking for another job, how they are looking for a job, how many jobs they have applied for these questions in order to secure as many valid answers as possible. Inconsistency checks were applied for these questions in order to secure as many valid answers as possible.

Health and Income (INC)
INC contains questions about health and about the income in 1992. "Block Health" is asked to all household members who are 16 years or older, and includes questions about height, weight, subjective judgements about one's own health, smoking and drinking habits, absence from work because of illness, and expected length of life.

Also "Block Income" was asked to all household members 16 years or older. Respondents were first asked about their personal income sources in 1992; that is, whether they were self-employed, received social benefits or pensions, had paid jobs on a contractual basis or any other income sources. Responses to these questions were used for routing. They were then asked to specify the yearly income they received from the relevant sources of income. They were also asked to specify transfers to children or other family members, the price of medical insurance, fringe benefits and inheritances and gifts.

The last questions in "Block Income" concerned household net income, and were asked to the heads of the households and their partners/spouses only. They were asked to estimate the household net income for 1992. They were also asked to specify what incomes they considered to be "very bad", "bad", "insufficient", "sufficient", "good", and "very good".

Information about income are thus sought in various ways. Questions are both on household level and individual level. Heads of households and their spouses/partners answered questions on household level. They were asked to give an estimate of household net income for 1992 using brackets. However, every respondent has also been asked more detailed questions about their gross incomes (fiscal profit for self-employed) from the various sources of income as well.

Note that not all household incomes are included in INC. Capital income is found in the WTH. Some questions about rent allowances and gifts are found in the HS.

INC also entailed a part in which respondents were asked to evaluate the questionnaire and their own responses. They were also asked to report how much time it took them to finish it. These questions are added to all questionnaires sent out by Stichting TelePanel. Some of the responses to these questions are reported below.

Accommodation and mortgages (HS)
This part of the questionnaire contains questions about current and former accommodation and about plans to move. It also contains questions about housing costs, like rent and mortgages. INC contains questions about health and about the income in 1992. "Block Health" is asked to respondents that currently have a paid job or used to have a paid job. The former group is asked about whether they have any additional work to the paid job they already reported, how many hours per week they work at this additional work and how many hours they would like to work. The latter group is asked about when they stopped working, for what reasons and how much they earned from their last job. Thereafter, both groups are asked about which pension plans and retirement insurance they take/took part in.

Assets (WTH)
WTH is the longest and most complicated part of the questionnaire with respect to routing. A large effort has been made in order to assure the value of assets as accurately as possible. Questions often begin with definitions of the specific assets under consideration in order to prevent misunderstandings of the questions. Responses to questions about balances can also be presented in a different way. First, respondents are asked to give the exact amount in guilders. Amounts can be given in brackets. This second option of giving answers results in more accurate data than the first, but less missing data. All household members 16 years or older answered questions about their private assets as of December 31st, 1992. With respect to shared assets, the respondents are told to report them only once per household. That is, if for
example the first member of the household reports shared assets, the other household members should ignore the questions about shared assets.

"Block Assets" contains initial questions about ownership of different kinds of checking accounts, saving accounts, saving certificates and other ways of holding assets. They are also asked about whether they plan to open new accounts in the next 12 months. They are then asked about ownership of cars, motorbikes, boats, caravans, real estate that is not serving as own accommodation, and debts to friends or family. The answers to these introductory questions are used for routing. There are separate blocks for each of the asset types. The respondents get the blocks that are relevant for them in which they are asked to report the balances of accounts and earned interest and they are asked to estimate the values of durables as of December 31st 1992. With respect to real estate, questions about any mortgages connected to them are also included. Note that not all assets are included here. Information about ownership of own accommodation is found in the HS described above.

"Block Debts/Liabilities" follows the same procedure as "Block Assets". The question block starts with questions that are used for routing. All household members 16 years or older are asked about the existence of private loans, extended lines of credit, outstanding debts with mail-order firms, loans from family or friends, study loans, debts through credit cards and other loans. Respondents are also asked about their opportunities to borrow money (liquidity constraints) and whether they would like to borrow money. Questions about mortgages on accommodation are not asked here. They are found in "Block Assets" (described above) for real estate which does not serve as accommodation for the household and in the HS for real estate which does serve as the household's accommodation.

Economic Psychology (PSY)
PSY starts with some questions that are to be answered by the head of the household and partner only. They are about income the past 12 months (note: another period than for the income measured in INC) and a subjective evaluation of this income, income development and expectations to income in the future (the next year and the next 5 years.) They are also asked about plans for saving, motives for saving, attitudes towards saving and debt, risk aversion, and changes in consumer expenditures. The questions about risk attitudes were only asked if the households' not income was above Dfl 20.000 over the past 12 months.

The next part of PSY contains questions which constitute scales developed to measure certain economic psychological variables like "time preference", "time horizon", "self control", "expected life time", "bequest motives", "self controlling strategies", "expectations of inheritance or gifts", and "personality". These questions are answered by all household members 16 years or older. Thereafter questions about family relations and perceived deprivation with respect to living conditions were asked. These were to be answered by the heads of the households and partners/spouses, and one more member of the household only.

Also the PSY entails questions where respondents are asked to evaluate the questionnaire and their own responses. These responses are discussed below.

2.3 Data Collection Schedule
Due to unforeseen problems, the first wave of data collection were delayed by 6 -7 months. The original plan was that all data should be collected in May-October 1993, when many people are occupied with preparing their tax returns for the previous calendar year. This might increase the validity of the data that are collected about financial matters. Due to the delay, data about 1992 were collected in December 1993 - April 1994 instead of in May-October 1993.

As the 2nd wave of data collection was carried out very soon after the 1st (in May 1994), it was neither recommended nor necessary to send out all the parts of the questionnaire also in the 2nd wave. It is possible that respondents would remember their responses to the questions in the 1st wave, and simply give the same answers to those questions in the 2nd wave. Moreover, it is reasonable to expect peoples health, living conditions and personal characteristics to be rather stable over the two months period between the 1st and 2nd wave. Therefore, only questions about assets and income were asked in the second wave (this time as of December 31, 1993 and income for 1993). However, as HS and WRK were not sent out in the second wave, mortgages on accommodation and the value of any owned accommodation are not measured as of December 31, 1993. This is unfortunate in cases where it is desirable to include changes in these variables in the definition of saving. Parts of PSY were sent out to the respondents who did not receive them in the first data collection due to a routing problem. All new households (replacement for panel members who dropped out) received all parts of the questionnaire.

An overview of the waves of data collection is found in table 2.2. Data will be collected in the coming years, as long as the project lasts.

<table>
<thead>
<tr>
<th>Period of Data Collection</th>
<th>Period in focus</th>
<th>Parts of the questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd wave</td>
<td>May - Oct. 1994</td>
<td>1993</td>
</tr>
<tr>
<td>3rd wave</td>
<td>May - Oct. 1995</td>
<td>1994</td>
</tr>
<tr>
<td>4th wave</td>
<td>May - 1996</td>
<td>1995</td>
</tr>
</tbody>
</table>

Table 2.2

Data Collection Schedule

2.4 Answering Questions about Financial Data
Many factors can influence the way people answer questions and thereby the accuracy or truthfulness of their responses. False responses, either given intentionally or unintentionally, can be caused by such factors as bad memory, misinterpretation of the question, self-image, and desires to answer according to perceived social expectations. The data-gathering approach might also have certain effects on the responses given. For example, differences can be found in anonymous versus identifiable self reports. Responses can also be affected by the ordering of
the questions, the length of the questionnaire, the level of threat perceived by the respondent posed by various types of questions, and question wording (Ferber, 1966; Sudman & Bradburn, 1974; Wentland, 1993). In this section we will focus on the data collection method and the subject of investigation; how people answer questions about financial matters. As no reliability tests for the responses from the STP and the HIP have been carried out yet, some results from other reliability tests are given below. They might give us an indication of the accuracy we might expect of the responses given to the Savings Project questionnaire.

According to Ferber (1966), the data reliability problem is more serious in savings surveys than in most other types of surveys due to two types of nonsampling errors; noncontacts (or refusals) and response errors. These errors produce biases in the estimates which tend to increase with sample size. In general, people who refused to participate in the surveys had higher savings than those who participated, and those who participated tended to under-report ownership of savings accounts. Also debt tends to be underestimated due to under-reporting of holdings of specific types of debt. The nonsampling errors vary substantially with type of financial holding. The more confidential an account or type of debt were perceived, the larger the errors. Ferber (1966) drew these conclusions based on a series of studies conducted to investigate the accuracy of savings statistics as obtained by consumer survey techniques. These studies were parts of the Consumer Savings Project of the Inter-University Committee for Research on Consumer Behaviour.

Based on a review of eight different debt-validation studies, Ferber (1966) reports that debt tends to be under-reported. Response error for reported debt was relatively small, and these response errors tended toward overstating rather than understating. However, the principal source of error in estimates of means or of totals was that respondents did not admit having a certain type of debt, and therefore no amount was reported. The more sensitive was the particular debt being questioned, the larger was the degree of understatement. For example, larger bias and larger nonsampling variances were obtained for personal loans than for automobile debt. In addition, a tendency to underestimate the frequency of small debt was found; and Ferber (1966) concludes from this that a large part of the underestimation of debt is due to bad memory.

Also with respect to time deposits, nonreporting of accounts was substantial, which lead to underestimation of holdings. The average size of the nonreported accounts exceeded by a substantial margin the average size of the accounts which were reported and for which balances were given. This was caused by two factors: 1) average holdings of households that could not be contacted or who refused to take part in the survey were higher than average holdings of households who were interviewed. In fact, refusals rate increased as the size of the accounts increased. 2) those who answered the questionnaire, failed to report the existence of accounts. Under-reporting was among other things associated with the interval between the date of the interview and the date of the last transaction to/from the account. However, when date of the interview and the date of the last transaction to/from the account. Ferber (1966) concludes that small amounts were not reported due to memory errors, while large amounts were not reported due to suspicion or fear.

Ferber (1966) found several factors which might influence reliability of data about savings and debt. First of all, the possibility of anonymity seemed to have an effect. Greater reliability was found when using a sealed-envelope technique by which respondents were offered the option of filling out the financial data themselves on a separate sheet, sealing it in an envelope addressed to the Survey Research Center and having the interviewer drop the envelope in the mail. The interviewer was requested in those interviews to show the respondent an IBM card and stress the anonymity of the data and the importance of the study. It was also found that mail questionnaires resulted in higher accuracy of reports than personal interviews, indicating that the presence of an interviewer might influence the accuracy of the data in a negative way. Secondly, it was found that more accurate estimates of saving were achieved when asking for balances on accounts at specific points in time, instead of asking for changes on an account over a certain period of time. Higher accuracy was also achieved when people were asked to consult their bankbooks before giving their answers, and use of an introductory letter before the interview took place, increased peoples willingness to do so. Third, shorter questionnaires resulted in higher reliability than longer ones, indicating that respondents lose patience after some time causing the accuracy in their reports to decrease. More accurate data were also achieved when the interval between the date of interview and date for which balances were requested, was short.

Wentland (1993) has reviewed and analysed data from 37 studies in which the validity of individual responses through using external criteria has been tested. The review covers a range of different subjects including financial matters, like tax evasion, salaries, bank accounts, and bankruptcy. Very different reports were given concerning the accuracy of the responses and response rates, and some of the reports about validity are not very encouraging. Note that the source of the reports given below is from Wentland (1993) and not the original sources.

Ferber et al. (1969a,b) carried out two validity checks specifically designed to check the validity of self-report data about savings account ownership and common stock holdings. The reference data used in the studies was about six or nine months prior to the interview date. Almost 50% of the respondents in their studies denied owning any savings account. Of the 421 respondents who reported ownership, 410 were willing to report the account size. Of these validating information was available for 399 respondents. Only about 10% of these reported their account size to the exact dollar, with over half of the distortions being understatements. Ferber et al. (1969 a,b) suggested that incompleteness of reports could be attributed to memory, as small accounts typically tended to go unreported. Problems with recall may also have affected accuracy on reports of the size of accounts, as they were asked about the balances on their accounts about 6-9 months in the past. Wentland (1993) suggests that inaccurate responses also might be caused by the fact that the key informants in these studies were the heads of the households. It is possible that this person did not know the balances of other household members' bank accounts to the exact amount.

Ferber et al. (1969a,b) found that reports about common stock holdings were more accurate than reports about savings accounts. 70% of the respondents reported ownership of stocks. Of those who were willing to report the number of common stock shares owned, almost 80% were truthful. Wentland (1993) argues that the more accurate reports with respect to common
stock holdings, might be due to greater stability in the number of shares owned versus the stability of dollar amounts in a savings accounts.

Maynes (1965) reported in a study of owners of savings accounts in the Netherlands Post Office Savings Bank that 95% of responses to questions about saving account ownership were accurate. However, he also reported that only 31% of the respondents gave accurate answer to the question about the balance on these saving accounts. The definition of accurate was, however, very strict, as deviations of more than one guilder were classified as "not accurate". In most of the cases the balances were under-reported. A tendency to over-report small balances and under-report large balances were found. Maynes (1965) suggests that the reason for this is that respondents wished to be perceived as average persons.

The results from Maynes' (1965) study were much more positive than those of Ferber et al. (1969a,b) with respect to accuracy. Wentland (1993) suggests that this is because all household members were interviewed about their personal accounts in the Maynes study and because respondents were asked about present balances and not about balance on the accounts 8-9 months earlier. Inaccuracy due to interviewing a "nonowner" and bad memory is thus avoided.

As a general conclusion, Wentland (1993) argues that among other factors, length of the recall period, the salience of the behaviour of which information is sought, time to reflect, and respondents' ability to check records or consult with other people, are potential determinants of accuracy. This indicates that memory is important for data reliability, and researchers should put more effort in helping respondents with remembering or letting them get a chance to consult records or other persons. Moreover, consistent with the conclusions drawn by Ferber et al. (1966), respondents' wish for social conformity, question sensitivity, and anonymity of the respondents are also important for data reliability. These factors influence both refusal rates and the tendency to either over- or under-report balances.

A recent study of the effect of interview method, confirms the importance of assisting the respondents in remembering. Grondil and Michaud (1994) set out to test whether there were any differences in quality on income data collected by paper and pencil (P&P) interviews and computer assisted interviews (CAI) respectively. Responses given in these interviews were compared with tax data for the same households. The computer questionnaire contained some interactive editing for certain types of income, something which was not possible for the paper and pencil questionnaire. The wages measured in the surveys differed from the tax data for 42.5% of the respondents to the paper and pencil interview and for 37.5% of the respondents to the computer assisted interview. For unemployment income (UI) the data were different for 17.2% (P&P) and 17.6% (CAI) of the respondents, while interest and dividends differed for 33% (P&P) and 23.2% (CAI) of the respondents.

Some of the differences between P&P and CAI might be explained by the fact that data were collected in different years. The paper and pencil interviews were carried out in 1991 while the computer assisted interviews were carried out in 1992. The tax data were from 1991. The unemployment rate was decreasing during these years, something that would result in an increase in people who report salary. However, the reports on UI benefits were quite stable, so it is difficult to assess the effect of the difference in survey year. In general, the survey results are closer to tax file data using CAI compared to P&P. There was also less under-reporting of UI and interest and dividends for the CAI. Grondil and Michaud (1994), attribute the better results from using CAI to the editing programme. In the CAI respondents were reminded about different types of income which they otherwise would forget. According to the authors, it seemed like under-reported income is due to lack of memory, or a failure of considering for example interest and dividends as income.

2.5 Results From the First Wave

In this section some results from the first data collection will be reported. Here, I will mention some of the technical problems which occurred in the 1st wave of data collection. I will also report some reactions from the respondents concerning the questionnaires and their own responses.

Technical Problems

The first data files received from Stichting TelePanel contained a large number of errors. They were mainly caused by problems with programming. Some households got the same questionnaires up to three times, something that caused double and triple occurrences of ID numbers of respondents. A wrong range check was applied for paid tax, so that reported tax above Dfl 31 were coded as missing (a range check for dates was used). These errors caused problems for people using the first data files that were available. However, these mistakes are now corrected, and the cleared files do not contain these errors.

Another problem which occurred in the 1st wave, was that the PSY was distributed in two parts (over two different weekends), and that the routing was not transferred from the first to the second part. This caused that a large number of respondents did not receive all the questions in the PSY which they were supposed to receive. This problem was solved by distributing the omitted questions on later occasions to the respondents who did not get them due to the routing error. Their responses to these questions are now included in the PSY data file, so that the number of missing cases for economic psychological variables, has been drastically reduced.

Respondents' reactions

For every round of data collection, Stichting TelePanel also includes questions where the respondents are asked to evaluate the questionnaires and their own responses. The files from the 1st wave of data collection contain some of the responses to these questions. Unfortunately, the responses to all these questions are not available.

The first evaluation questions the respondents got are shown in the table 2.3 below. The answers to these questions are found in the PSY and the INC files. However, these responses do not only concern the PSY and the INC. As the PSY was distributed in two parts, the evaluation will only concern one of these parts of the PSY. Likewise, the evaluation given by the HIP members concerns more questions or parts of the questionnaires than the evaluation given by the STP members, as they receive bigger portions of questions every time.
Means and medians for responses from the full sample are also reported in the table. If we compare means and medians for the answers of the STP members and the HIP members respectively, the STP members are in general more positive in their evaluation of the PSY. The difference is largest for the questions concerning to what degree the questions could hold their interest and how much time it took them to fill out the questionnaire. With respect to the INC, the HIP members are more positive than the STP members. The differences are largest for the question concerning how easy the questions were to answer and the question about how long it took to fill out the questionnaire. The differences with respect to evaluation of time spent on the questionnaire, are probably due to the fact that HIP-members and STP members receive questionnaires of different length.

Finally, we would like to ask your opinion on the questionnaire that you have answered just now. Below, you can assign 1 up to 10 points to a number of aspects. 1 = very bad, 10 = excellent. If you really do not know, type a 0 (zero).

How many points would you assign to the QUESTIONNAIRE with respect to:

| - the degree to which the questions could hold your interest | INC: mean= 6.0 and median= 6.0 |
| - how easy the questions were to answer | PSY: mean= 6.2 and median= 7.0 |
| - how clear the questions were | INC: mean= 6.7 and median= 7.0 |
| - the lay-out of the questions as shown on your screen | PSY: mean= 6.9 and median= 7.0 |
| - how long it took you to fill in the questionnaire | INC: mean= 7.1 and median= 7.0 |
| | PSY: mean= 7.0 and median= 7.0 |
| | INC: mean= 7.4 and median= 8.0 |
| | PSY: mean= 7.2 and median= 7.0 |
| | INC: mean= 6.6 and median= 7.0 |
| | PSY: mean= 5.8 and median= 6.0 |

Table 2.3
Respondents' Evaluation of Parts of the Questionnaire

In general, respondents were positive about the questionnaire about health and income (INC), as most of the ratings given were from the positive side of the scale. 70.4% of the respondents rated the degree the questions could hold their interest at 6 or better. 79.3% of the respondents answered 6 or better to the question about how easy the questions were to answer. 88.4% rated 6 or better to the clarity of the questions. 93.2% rated the lay-out of the questionnaire to six or better. Finally, 81.2% evaluated the length of the questionnaire (tijdsduur) to 6 or better.

Also the evaluation of the PSY was overall positive. 71.4% of the respondents rated the degree the questions could hold their interest to 6 or better. 85.1% of the respondents rated 6 or better to the question about how easy the questions were to answer. 87.1% rated 6 or better to the clarity of the questions. 89.9% rated the lay-out of the questionnaire to six or better and 62.9% evaluated the length of the questionnaire (tijdsduur) to 6 or better.

The next evaluation question goes as follows (percent in parentheses):

What would you say about the answers you have given in the questionnaire:
1 (almost) all of them are correct (56.4%)
2 most of them are correct (39.6%)
3 most of them are wrong (4.2%)
4 (almost) all of them are wrong (6.7%)
If you really do not know, type a 0 (zero)

Responses to this question is only available for the INC. 95.4% of the respondents answered that almost all (1) or most (2) answers are correct. 1.9% said that most (3) or almost all (4) of the questions were wrong. This corresponds to 92 respondents. These 92 respondents received the following follow up-questions (frequencies in parentheses):

Please indicate why you think (almost) all of your answers are wrong:
More than one answer is possible here:
1 the questionnaire did not match with my situation (22)
2 there was not enough opportunity to correct mistakes that I had made (15)
3 my answers were wrongly recorded by the computer (1)
4 the way the questioners were put was too complicated (8)
5 it was too difficult to remember the things that were asked, and/or it was too much work to look them up (41)
6 other reason (38)

38 respondents reported that they had other reasons for thinking that nearly all their responses were incorrect. Some of these households answered that they had not discovered that the questions were about income in 1992. They had reported income for 1993 instead.

The next evaluation question was:
How many times did you make use of the option to correct answers you had already given? You can type a number from 1 to 5, where 1 = NEVER and 5 = VERY OFTEN.
If you really do not know, type a 0 (zero)

86.1% of the respondents answered 1 or 2 to this question, indicating that they never or seldom used the option to correct answers.

In the last evaluation question, the respondents were asked to comment upon the questionnaire:
Do you have any remarks about the questionnaire you have answered just now?
1 no
2 yes
You can use the rest of the screen for your comments.

936 of the respondents answered "yes" to the question about they had comments about the INC, and 484 respondents answered that they had comments about the PSY. A qualitative analysis of these comments could be interesting and might provide useful feedback concerning formulations of the questions and about how sensitive the questionnaires were perceived by the respondents. So far, such an analyses has not been conducted.
2.6 Summary and Conclusions

Data for the VSB-CentER savings project are collected by Stichting TelePanel, University of Amsterdam, using Computer Assisted Panel Research (CAPAR). The sample consists of two sub-samples; one sample intended to be representative of the Dutch population (containing about 2000 households) and one sample intended to be representative of Dutch Households with the 10 percent highest incomes of the Netherlands (containing about 1000 households). Data for the savings project will be collected in May every year as long as the project lasts. So far, four waves of data collection have taken place, covering income and saving for 1993, 1994 and 1995 and holdings as of December 31st, 1992, 1993, 1994 and 1995.

The questionnaire is divided into five parts covering different aspects of economic behaviour as well as characteristics of the households. All household members above 15 years were asked to answer the questions. As it has been stressed that the person who is most likely to know the correct answer to the different questions should be interviewed, most questions are given to all household members (excluding those below 15). For example, it is assumed that the probability of getting accurate responses is highest when the owner of specific assets is asked about their value. Every household member therefore reports about values and balances of their own assets. When information about the total household or about shared assets is sought, the household members considered to be the most knowledgeable about these matters, are used as key informants.

Advantages and disadvantages associated with CAPAR have been outlined. Computerised interviewing allows complex routing and high control over question flows. This makes tailoring of the questionnaires for different respondents possible, reducing the workload of the interviewers. Interviewer effects are also avoided. The disadvantages associated with CAPAR are that the routing might cause a chain reaction of mistakes, if the respondent gives a wrong answer to one of the questions used for routing. In some cases the routing also becomes too complex for the programmer to handle, something that also might lead to missing data. Panel conditioning might cause that the sample over time becomes atypical for the population it is supposed to represent.

Results from previous studies of saving behaviour indicate that it is reasonable to expect both savings and debt to go under-reported. It will therefore be difficult to produce aggregate estimates of savings and debts for the Dutch population based on the panel data. Validity checks will also be difficult to carry out in order to assess the magnitude of the underestimates. The reason for this is that it is not the reported balances or amounts that tend to be inaccurate, but the reports of ownership of certain assets or holdings of specific types of debt. A check of the accuracy of the reports will therefore require access to customer registers of all financial institutions in the Netherlands, as we would have to search for the unreported accounts.

However, there is no reason to expect that the VSB data are less reliable than data collected for other savings projects by other data collection methods. In previous studies, it has been found that bad memory is one important source of low reliability of financial data. When constructing the questionnaire for the Savings Project, substantial effort has been put in explaining the respondents what kind of information that is sought by defining what is meant by different types of savings accounts, mortgages, assets, etc. All different waves of saving or holding debt are also mentioned, something which might help respondents in remembering. As there is no interviewer present, the respondents have time to consult records and other household members without feeling any time pressure. This might also increase data accuracy.

In addition, the different range checks and confirming questions which are included in the programme, will also contribute to higher accuracy. Moreover, as data are collected from the same respondents several times, it is possible to check data based on data collected in subsequent waves. Finally, the reactions from the respondents to the questionnaire are generally positive, something that also gives an indication that respondents have tried to give accurate responses.

As the 1st wave was delayed, this caused the time interval between the date of the interview and the data of the balances which were to be reported to be more than a year. This might have a negative effect on the data accuracy. It is therefore reasonable to expect higher accuracy for the data collected for the 2nd and 3rd wave, for which the intervals between the dates of the interviews and the period for which informations sought, are shorter.

It could also be the case that data collected from the STP are more reliable than data collected from the HIP. Previous investigations show that the length of the questionnaire has a negative relationship with data accuracy, as the respondents get tired from answering questions. The HIP members do not only get the questions in bigger portions, it is also likely that the HIP members will get routed to more questions than the STP members, as they might have more assets and incomes to report.

The panel conditioning might be a threat to both internal and external validity of studies of causes and effects using the VSB panels. Panel members might change their behaviour due to the panel membership, and this might be a competing explanation of observed behaviour to other proposed causes. The observed behaviour in the panel might therefore be atypical for the populations they are supposed to represent, and this limits the possibilities of generalising results obtained from analyses of the VSB data.
3. Sampling Procedures

The two panels and the sampling schemes applied when recruiting respondents for the panels are described in detail below. Possibilities of sampling errors are discussed as well as biases due to nonsampling errors (effects of nonresponse).

3.1 The STP

The STP is intended to be representative of the Dutch population with respect to socio-economic variables like region, political preference, housing, income, degree of urbanisation, and age of the head of household. This panel is a part of the ordinary panel used by Stichting TelePanel (the TelePanel Foundation). Hence, the members of this panel are not recruited specifically for the VSB project, but they are members of a permanent panel and serve as informants for a variety of research projects. The STP was set up in 1991, which means that the members of this panel have participated in the panel for 5 years or less.

The sampling was carried out by using telephone directories as sampling frame. Drawn telephone numbers were used as basis for creating so-called 100-banks of telephone numbers. This was done by randomising the two last digits of every drawn number. Telephone numbers were thereafter drawn from the 100-banks. In this way, also new lines and nonlisted numbers had a chance of being selected. The goal for the sampling was to recruit 2000 households for the panel. In order to achieve this, Stichting TelePanel drew 12153 numbers, based on expected responses. Expected responses for the different stages of the recruitment process were as follows:

It was expected that 57% of the households who in telephone interviews said they were willing to co-operate in the panel, would be selected for participation by Stichting TelePanel. That means that 3500 households would have to be willing to participate in the STP in order to recruit 2000 households. These 3500 households would have to be approached and introduced to the computerised interviewing technique by representatives from the Stichting TelePanel.

It was expected that 80% of the approached households would be willing to co-operate. This means that 4375 households had to be approached by telephone interviews in order to find 3500 households who were willing to co-operate.

The response rate for the telephone interviews was expected to be 60%. This means that 7292 households had to be called and asked whether they were willing to participate in the telephone interview. Finally, since only 60% of the drawn numbers were expected to be residential, 12153 numbers would have to be drawn.

‡ Unfortunately, the actual response rates for the sampling of the STP are not available. The response rates (and corresponding numbers of households) reported here, are based on response rates for the recruitment process when additional sampling was carried out in 1992.

‡ This number seems to be too low, as according to CBS statistics, there were 4.3 million households in The Netherlands in 1992 (at least 95% of these households have a telephone). In addition, a large number of companies and institutions have telephone.
representative with respect to degree of urbanisation and region. In general, a community was defined as self-representing if it contained more than \(17.4 \times 434 = 7552\) telephone numbers (that is: the expected number of telephone numbers which had to be drawn in order to give five panel members \(\times\) the sample fraction). In other words, if a community contained so many numbers that 5 or more numbers could be drawn with a probability of \(1/434\), it was sampled with certainty. According to this rule, 158 communities were self-representing. Together, these communities contained 3,220,000 numbers and should deliver 3,220,000/434 = 7,442 numbers.

The 545 non-self-representing communities, which contain 2,050,000 numbers, had to deliver 12153 - 7,442 = 4,711 numbers. The small communities were selected with probabilities proportional to the size of the population living in the communities. For this stage, 4711/17.4 = 271 small communities were drawn.

The sampling fraction \(f\) for the selection of telephone numbers from the small communities is found by multiplying the sampling fraction for drawing a telephone number in the first stage \(f_1\) with the sampling fraction for drawing a telephone number in the second stage \(f_2\), so that \(f = f_1 \times f_2\). In this way, the probability for each telephone number to be drawn is held constant. For the telephone numbers in small communities, this can be expressed as

\[
f = f_1 \times f_2 = \left( \frac{271 \times B_i}{20,500,000} \right) \times \left( \frac{17.4}{B} \right) = \frac{1}{434}
\]

where \(B_i\) denotes numbers of telephone numbers in cluster \(i\). Hence, the probability for a telephone number to be drawn is 1/434 both in self-representing and drawn communities.

Selection of Secondary Sampling Units

Telephone numbers drawn within the selected clusters were used to create so-called «100-banks» of telephone numbers. The PTT (Post Telefoon Telegraaf) -phone files were used as a database for the sampling. Within the selected clusters 12,153 telephone numbers were drawn. The sub-sampling rate within each cluster had to vary in order to keep the sampling fraction constant according to the procedures outlined above. For the big communities the numbers of telephone numbers drawn within each cluster was proportional to the cluster’s size. For the smaller communities a fixed number of telephone numbers was drawn. If, for example, the number 0245476267 was drawn, a 100-bank was created by randomising the last two digits in this number. This procedure is called List Assisted Random Digit Dialling (LARDD) or Directory Assisted Selections. LARDD has the advantage that both secret numbers and new lines can be drawn and included in the sample. The disadvantage is that non-residential numbers (nonexisting numbers or companies) will be included.

Selection of telephone numbers

12,153 telephone numbers were selected within the 100-banks. One number was selected from each 100-bank. The fraction of non-residential numbers was lower than expected. 80% of the drawn numbers turned out to be residential which means that 9,722 eligible telephone numbers were dialled. The low fraction of non-residential numbers was probably due to the fact that the LARDD was used, and not the traditional random digit dialling. In USA, residential telephone numbers are often clustered within banks of telephone numbers. Use of listed residential numbers as a basis for randomisation will thus increase the probability of hitting residential numbers (Lepkowski, 1992). Probably, this is also the case in The Netherlands.

Selection of households

The nonresponse for the next stages in the recruitment process was higher than expected. The households were first contacted by phone. In the first interview, the respondents were asked about background information and whether they were willing to participate in the STP. The response rate for the first recruitment interviews was about 70%, which means that 6,806 households participated. About 50% of the households who took part in the recruitment interviews were prepared to participate in the STP. These 3,403 households were visited by representatives from Stichting TelePanel and introduced to the computer aided interviewing. About 1,804 of the households (53%) who were introduced to the computer aided interviewing technique participated in the panel eventually. The overall response rate for the STP is thus 18.5% (1,804 panel members of 9,722 eligible telephone numbers).

Random sampling means that the numbers that are to be included in the sample are selected at random and that each unit has an equal chance of inclusion in the sample. This was the strategy used for selecting telephone numbers. However, if there is nonresponse among the sample cases, the unbiasedness property which random sampling is assumed to provide, does not apply. Nonresponse makes it impossible to obtain complete measurements on the survey sample. As shown above, the level of nonresponse when recruiting members for the STP panel was high (81.5%). Table 3.1 below also shows that people who were willing to participate were different from those who refused. This means that if all households willing to participate in the panel were included in the sample, the selective nonresponse would lead to a sample which was not representative of the Dutch population. This is demonstrated for the variables "composition of household" and "age" in table 3.1 below.

<table>
<thead>
<tr>
<th>Table 3.1</th>
<th>Distribution of selected households that eventually joined the panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution of respondents selected to join the panel (50.4% of the interviewed households)</td>
<td>Distribution of respondents selected to join the panel (50.4% of the interviewed households)</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
<tr>
<td>Singles</td>
<td>27.0</td>
</tr>
<tr>
<td>Family with children</td>
<td>33.2</td>
</tr>
<tr>
<td>Family without children</td>
<td>39.8</td>
</tr>
<tr>
<td>&lt; 29 years</td>
<td>18.9</td>
</tr>
<tr>
<td>30 - 39 years</td>
<td>21.5</td>
</tr>
<tr>
<td>40 - 49 years</td>
<td>18.3</td>
</tr>
<tr>
<td>50 - 64 years</td>
<td>21.4</td>
</tr>
<tr>
<td>&gt; 65 years</td>
<td>19.2</td>
</tr>
</tbody>
</table>

Source: Stichting TelePanel. The figures are from an additional sample done to compensate for panel attrition. 4,596 households were interviewed by telephone, and 2,813 households were willing to participate in the panel.

Table 3.1

Responses for Different Stages in the Recruitment Process.
Table 3.1 shows that households with older members, one-person households and families without children were less willing to participate in the panel than others. In order to correct for this non-random error, Stichting TelePanel assigned quotas in an unbiased way for some background data. Quota sampling means that the mix of types of households who can enter the sample, is controlled. For the STP, the selection was carried out so that the sample would be representative with respect to region, political preference, housing, degree of urbanisation, income, and age of head of household. As the sample is made up by selecting households who it is presumed will give the sample the same characteristics as the population it should represent, this can also be called purposive selection (Neyman, 1934). The rather low response rate in the last stage of the recruitment process is thus caused by two factors: refusals to cooperate and nonselection by Stichting TelePanel.

As the sampling procedure described above did not result in as many panel members as expected (1803 instead of 2000 households), additional sampling had to be carried out. This was done by drawing more numbers from the 100-banks which were created during the first sampling. In addition, new 100-banks were created following the same procedure as described above. In the end, a panel consisting of 2188 households was established. When the 1st wave of data collection for the VSB-CentER Savings Project took place, the STP consisted of 4462 persons in 1880 households.

As the STP suffers from a 30% annual attrition rate, Stichting TelePanel has to perform continuous surveys in order to bring in new panel members. This is done by drawing numbers from the created 100-banks. So far, four or five numbers from each of the 100-banks have been drawn for this purpose. The telephone interviews of selected households are carried out, which results in a reservoir of households willing to cooperate in the panel. Households who drop out are replaced every week with households with similar characteristics using purposive selection as described above. If some of the households split, Stichting TelePanel tries to follow both partners in the following periods.

Table 3.2 gives an overview of distributions of some of the variables that are taken into consideration in order to make STP as representative of the Dutch population as possible: degree of urbanisation, age, income, composition of household, and housing situation. Distributions for these variables are shown for the Dutch population above 16 years (when such data are available) and for the panel members that eventually responded to the so-called PSY questionnaire (see section 4). In addition we compare distributions of occupation and education between the STP members and the HIP members.

3.2 The HIP

The members of the HIP were recruited specifically for the VSB project because of a wish to study high-income households in more detail. As income and wealth distributions in populations often are skewed, a simple random sample would not provide enough wealthy households to draw conclusions about their behaviour and distributions of many financial variables (Kennickell & McManus, 1993). The high-income households are defined as households with the 10% highest incomes of the Netherlands. These households have a gross income of at least 105,000 guilders per year. According to Stichting TelePanel, the sample is representative of this part of the Dutch population with respect to income, age, gender and degree of urbanisation.

The goal for the sampling was to recruit 1,000 households to the panel. A problem related to this recruitment process was that the target population was much smaller than the sample population (households with a telephone). Hence, the probability for drawing a number for a high-income household was much smaller than for hitting a household that would fit for the STP. If the sampling for the HIP was to be done in the same way as for the STP, 100,000 numbers would have to be drawn and dialled based on the same expected response rates.

Drawing and dialling 100,000 numbers would be an expensive and exhausting task. In order to reduce costs by reducing the number of households that would have to be approached, two different sampling schemes were used for selecting households to the HIP. They differed in method and in hitting chance for the target population.

The first procedure was as follows:

1) Selection of area post codes for wealthy areas
2) Selection of telephone numbers within the selected wealthy areas
3) Selection of households

Selection of area post codes

A database belonging to a Dutch company called «Geo-Marktprofile» was used in order to increase the probability of approaching households belonging to the target population. Their database consists of 400,000 postal codes areas. Geo-Marktprofile constructed a variable called «wealth class» based on variables like prices of houses (also rent), presence of gardens and level of prosperity. In this way it was possible to get an overview of the geographical distribution of Dutch households over wealth categories.

Stichting TelePanel has information about distributions of household income where the households are obtained via an unbiased sample (the STP). These two sources of information were matched so it was possible to determine the distribution of high-income households over the wealth categories. Based on this information, it was possible to estimate how the 600,000 high-income households in the Netherlands are distributed over wealth categories. The areas were categorised within 6 different classes of prosperity, where “1” represents the most wealthy
area. The areas which were classified as 1 ("high") or 2 ("above average") were selected. These areas covered about 55% of the target population.

Selection of telephone numbers in wealthy areas

Households were selected within the wealthy areas. Stichting TelePanel stated how many phone numbers they needed from each wealth class and the numbers were drawn by Geo-Marktprofiel as they had the database with all addresses and listed telephone numbers. Use of Geo-Marktprofiel's database increased the probability of approaching a high-income household from 10% to 20%. In addition, all the numbers which were drawn were residential. The disadvantage with this sampling scheme is that it does not include new and unlisted telephone numbers. As 12% of all telephone numbers in the Netherlands are not listed, this might cause a coverage error in the sample. 20,000 households were approached based on this sampling scheme.

In addition, another 10,000 households were approached. These households were selected using the same 4-step sampling scheme as described for the STP. This scheme covers nearly 100% of the households in the target population. The probability for hitting a high-income household using this method, was 10%.

Selection of households

In total, 30,000 households were approached. In the first interview they were asked about some characteristics of the household (among them income). About 66% of the respondents answered the questions about income. Households belonging to the target population were asked whether they were willing to participate in the telepanel. These households were informed about the purpose of the study they were asked to take part and why they were approached. 50% of the households in the target income group were willing to participate in the panel. More than half of the households who said they were willing to become panel members, participated in the high income panel eventually. In the end, 1011 households participated in the panel.

The hitting rate of a rich household was 20% for the 20,000 household approached using the first sampling scheme, while it was 10% for the 10,000 households approached using the second sampling scheme. This means that about 5,000 households in the target income group were approached. As 1.011 of these households participated in the panel, this corresponds to a response rate of 20.2%. When data collection for the 1st wave of VSB-CentER Savings Project took place, the HIP consisted of 2,735 persons in 910 households. An overview of the response rates for the 1st wave is given in table 3.3. Unfortunately, more accurate information about the sampling and response rates is not available.

Also the HIP suffers from attrition, but the rate of attrition is not known. Selections of households when replacing the households which drop out, is done by using the 100-banks created for the STP.

In table 3.2 distributions for socio-economic variables like income, age, gender, composition of household, degree of urbanisation, education, housing situation and occupation for the STP and the HIP members are shown.

---

1 Definitions compared to CBS-statistics: Large towns = zeer sterk stedelijk, small rural = sterk stedelijk, and rural area = niet stedelijk
2 Codels of education: Low education = Kleine/medium groot/medium, special education, VWO special education, Low middle education = LBO, HBS, VGS, MAVO, LAWO and kliniek wonen, High middle education = HAVO, VWO and MBO, High education = Hoger beroepsopleiding, wiskunde/lopende programmering
3.3 Distributions of Some Socio-Economic Variables

Table 3.2 shows the distribution of some socio-economic variables for the Dutch population, the STP, and the HIP. According to their figures, the HIP is not representative with respect to composition of the household. Differences between the STP and the HIP. The HIP and STP members differ with respect to other variables than income. These are variables which typically are related to income. A more detailed description of the two panels and the differences between them can be found in the technical report by van Dijk and Beckman (1994). Below, only a few of the differences will be reported.

More HIP members had at least one job in 1992 than STP members. There are also more self-employed people in the HIP than in the STP (van Dijk & Beckman, 1994). This can be seen in table 3.2, as more people in the HIP reported to have a paid job. STP members receive social benefits, retirement benefits, and illness and unemployment benefits more often than HIP members, while HIP members have income from real estate more often than STP members (van Dijk & Beckman, 1994).

Females in the HIP are more often employed than females in the STP. This means that a higher fraction of the households in the HIP than in the STP consists of households with double income (two adults with income). Income per person in the HIP is therefore not necessarily very high, it is the participation of women in the labour market that causes the relatively high household income.

The distribution of degree of urbanisation differs between Dutch population and the STP. However, this might be due to different coding by Stichting TelePanel, and it is therefore difficult to conclude anything about representativeness with respect to urbanisation.

The HIP

As I have not succeeded in getting statistics for "Dutch households with the 10 percent highest incomes of the Netherlands" from the CBS, it is not possible to check to what degree the HIP is representative of the population it is drawn from. The only data available with respect to this special part of the Dutch population are found in the report written by van Dijk and Beckman (1994). Their source is Stichting TelePanel (see "income" and "composition of household" in table 3.2). According to their figures, the HIP is not representative with respect to composition of the household.

<table>
<thead>
<tr>
<th>Occupation - Head of household</th>
<th>STP-members who responded to the PSY - 1. wave %</th>
<th>HIP-members who responded to the PSY - 1. wave %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paid job</td>
<td>57.9</td>
<td>75.8</td>
</tr>
<tr>
<td>Unpaid job</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Unemployed</td>
<td>3.2</td>
<td>0.9</td>
</tr>
<tr>
<td>Military service</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Student</td>
<td>2.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Working in own household</td>
<td>6.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Retired (AOW + VUT)</td>
<td>20.2</td>
<td>10.9</td>
</tr>
<tr>
<td>(Partly disabled)</td>
<td>4.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Other</td>
<td>3.6</td>
<td>4.6</td>
</tr>
</tbody>
</table>

| Occupation - all respondents   | Medicaid
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paid job</td>
<td>47.5</td>
</tr>
<tr>
<td>Unpaid job</td>
<td>1.6</td>
</tr>
<tr>
<td>Unemployed</td>
<td>2.5</td>
</tr>
<tr>
<td>Military service</td>
<td>0.0</td>
</tr>
<tr>
<td>Student</td>
<td>7.5</td>
</tr>
<tr>
<td>Working in own household</td>
<td>20.9</td>
</tr>
<tr>
<td>Retired (AOW + VUT)</td>
<td>13.3</td>
</tr>
<tr>
<td>(Partly disabled)</td>
<td>3.2</td>
</tr>
<tr>
<td>Other</td>
<td>3.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Net Income per month</th>
<th>For hh. with the 10% highest income</th>
<th>Full panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 6660</td>
<td>0.0</td>
<td>44%</td>
</tr>
<tr>
<td>6661 - 8.910</td>
<td>42%</td>
<td>37%</td>
</tr>
<tr>
<td>8.911 - 12.500</td>
<td>15%</td>
<td>16%</td>
</tr>
<tr>
<td>&gt; 12.501</td>
<td>4%</td>
<td>3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Net Income per month</th>
<th>Full panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 1.700</td>
<td>18%</td>
</tr>
<tr>
<td>1.701 - 2.500</td>
<td>19%</td>
</tr>
<tr>
<td>2.501 - 3.300</td>
<td>19%</td>
</tr>
<tr>
<td>3.301 - 4.200</td>
<td>20%</td>
</tr>
<tr>
<td>4.201 - 5.000</td>
<td>23%</td>
</tr>
<tr>
<td>&gt; 5.001</td>
<td>9%</td>
</tr>
</tbody>
</table>

2) Centraal Bureau voor de Statistiek

Table 3.2
Overview of Distribution of Background Variables for the Dutch Population, the STP, and the HIP

3.3 Distributions of Some Socio-Economic Variables

Table 3.2 shows the distribution of some socio-economic variables for the Dutch population and the respondents to the PSY questionnaire (1st wave). It is not easy to check whether the two panels are representative of the populations they are drawn from. First of all, it is difficult to find distributions for the Dutch population for persons younger than 16 years are excluded. Most official statistics also include people below 16 years. Direct comparisons with the distributions of the panel members are therefore not possible. It is also difficult to find distributions of households with the 10% highest incomes in the Netherlands and distributions for heads of households only. This means that we have to rely on Stichting TelePanel when they say that the panels are representative of the populations they are drawn from with respect to certain socio-economic variables (like for income, composition of the household, and degree of urbanisation (see table 3.2)). Moreover, comparisons between distributions in the Dutch populations and the panels are only possible for some variables. Although the STP might be representative of the Dutch population with respect to certain socio-economic variables, this might not be true for behavioural and psychological variables.

The STP

Table 3.2 shows that women, old and young persons, and tenants are under-represented in the STP when we compare the distributions among respondents with CBS statistics. According to Stichting TelePanel, the samples are representative with respect to income and degree of urbanisation.

The distribution of degree of urbanisation differ between the Dutch population and the STP. However, this might be due to different coding by Stichting TelePanel and CBS, and it is therefore difficult to conclude anything about representativeness with respect to urbanisation.

The HIP

As I have not succeeded in getting statistics for "Dutch households with the 10 percent highest incomes of the Netherlands" from the CBS, it is not possible to check to what degree the HIP is representative of the population it is drawn from. The only data available with respect to this special part of the Dutch population are found in the report written by van Dijk and Beckman (1994). Their source is Stichting TelePanel (see "income" and "composition of household" in table 3.2). According to their figures, the HIP is not representative with respect to composition of the household.

Differences between the STP and the HIP

The HIP and STP members differ with respect to other variables than income. These are variables which typically are related to income. A more detailed description of the two panels and the differences between them can be found in the technical report by van Dijk and Beckman (1994). Below, only a few of the differences will be reported.

More HIP members had at least one job in 1992 than STP members. There are also more self-employed people in the HIP than in the STP (van Dijk & Beckman, 1994). This can be seen in table 3.2, as more people in the HIP reported to have a paid job. STP members receive social benefits, retirement benefits, and illness and unemployment benefits more often than HIP members, while HIP members have income from real estate more often than STP members (van Dijk & Beckman, 1994).

Females in the HIP are more often employed than females in the STP. This means that a higher fraction of the households in the HIP than in the STP consists of households with double income (two adults with income). Income per person in the HIP is therefore not necessarily very high, it is the participation of women in the labour market that causes the relatively high household income.
The higher participation of women in the labour market in the HIP is probably related to the level of education. More people in the HIP have a high education than in the STP, and more people left education with a diploma. 55.7% of the members of HIP has a high education, while the corresponding number for the STP is 28.1%. If we only look at the heads of the households, the difference between the panels is even clearer: 74.6% of the heads of the HIP have a high education. while holds. the difference between the panels is even clearer: 74.6% of the heads of the HIP members more often own. their own house than the STP members. In the HIP, as many as 91.4% of the households own their accommodation, while the corresponding figure for the STP is 59.6%. According to van Dijk and Beekman (1994), HIP members live less often in a flat or a single-family terraced house, as they have a higher tendency to live in (semi-)detached single-family houses.

People in the HIP have in general more assets than people in the STP and have more opportunities to save and invest money. HIP members use saving accounts and other forms of saving or investing more often than STP members - especially complex forms of saving and investing like single premium insurance and investments in stocks (van Dijk & Beekman, 1994).

HIP members perceive their health to be better than the STP members do. More people in the HIP suffer from long-term illnesses and handicaps. HIP members also smoke less (van Dijk & Beekman, 1994).

3.4 Sample Biases
The STP is meant to be the Dutch population in miniature. This goal has also been reflected in the sampling procedures applied when the panel members were recruited. Random selection of telephone numbers have been combined with stratification and quota sampling. Table 3.2 shows that the STP is representative of the Dutch population with respect to some socio-economic variables. The HIP is also intended to be representative of Dutch households with the 10 percent highest incomes of the Netherlands. However, as shown above, it is difficult to judge whether the samples are representative as is it is difficult to obtain comparable information about the target populations.

Below some possible sampling biases caused by the sampling frames applied and the recruitment procedures, will be discussed. These biases might represent possible threats to validity in studies of causes and effects with respect to economic behaviour. They might therefore limit the extent to which results from analyses of the VSB-data can be generalised across populations, situations and times. This does not mean that each bias operates with equal frequency or affects all results to the same degree. However, researchers using the data should be aware of them and evaluate their impact in each individual study.

Bias caused by the sampling frame applied
Interviewing using modems requires that the respondents have a telephone. As the panel members also are recruited by telephone interviewing, none telephone households are not included in the sampling frame. The potential for coverage errors in telephone surveys decreases as the percentage of telephone households increases. According to Sichting TelePanel, 97% of the Dutch households have a telephone, so the sampling frame should therefore not represent a serious threat to representativeness. It covers nearly the total target population. However, if there are large differences between telephone and non telephone household populations, there still might be a chance of a significant noncoverage bias.

According to Trevin and Lee (1988), noncoverage in telephone surveys in the Netherlands is higher for single and divorced persons, households where the head is not skilled, households with an unemployed head, households with a young head, large households (6+ persons), and it is higher for households living in rental accommodation than owner-occupied accommodation. Relatively to other countries, the Netherlands exhibits little differences in telephone coverage across income groups, but a tendency for noncoverage for low income households can still be found (Trevin & Lee, 1988).

Also CBS-studies show that the young, the elderly, the unemployed, divorced people, people with lower education levels and people in lower income groups less frequently have a telephone than others (Snijders, 1992). The sampling frame used for the sampling, might therefore be the reason why tenants and young and old people are under-represented in the STP. The number of students and unemployed also appears to be low in the STP, although comparable data for the Dutch population are lacking (see table 3.2). The applied sampling frame also excludes people living in institutions of various kinds, those in military service and transient.

The sampling frame might cause certain types of households to be overrepresented. 1.5% of the Dutch households have more than one telephone number, so they get a higher probability for being selected than households with one telephone number. Having two telephone lines is often related to being self-employed and having high income. This part of the population might therefore be overrepresented in the STP.

The most important sampling frame used for recruiting members for the HIP, covered only 55% of the target population. These were high-income households living in areas which can be categorised as being wealthy (because of high prices of houses, high rents, and presence of gardens). This means that high-income households who choose not to spend so much money on accommodation and therefore live in less prosperous areas, have a relatively low probability of being included in the sample. They had a chance of selection due to the second sampling frame applied (which includes all households with a telephone), but their chance of selection is much lower than for households which live in wealthy areas. The HIP panel is therefore biased with respect to the priorities the households make when spending and investing their relatively high income. In this respect, the HIP is not representative of the target population.

Bias caused by refusal to co-operate
The samples suffer from a high level of nonresponse. As mentioned earlier, the anti-bias property that random sampling is supposed to provide, does not necessarily apply when we have nonresponse. As shown in table 3.3, below, nonresponse was high, which increases the probability of nonresponse errors.
Table 3.3 shows the response rates from the 1st wave of data collection for the STP and the HIP respectively. The response rate \( (A \times B) \) is defined as the number of completed interviews divided by the total number of eligible sample units. The response rates reported here are on household level. The response rates for panel participation \( (A) \) are based on information from the additional sampling carried out in 1992 for the STP, and on certain assumptions with respect to hitting rates of residential households in the target income group for the HIP. Table 3.3 shows that the response rates for panel participation are relatively low. The response rates reported in column \( A \) are a product of two different mechanisms: 1) the households who were approached but declined to become panel members (refusals) and 2) households who were willing to participate in the panel but were not included in the sample due to the quota sampling (nonselection). As it is not possible to separate these mechanisms' relative contribution to nonresponse, the response rates in column \( A \) are a product of both types.

In column \( B \), questionnaire-specific nonresponse is reported. This is nonresponse caused by panel members who for some reason did not fill in the different parts of the Savings Project questionnaire (here: the HS). The response rates within the panels were relatively high; 94.8% for the STP while it was 99.4% for the HIP (on household level). The overall response rates (column \( A \times B \)) are a product of both refusals to participate and questionnaire-specific nonresponse.

Finally, we have item nonresponse, which is failure to answer specific questions in the questionnaire. For example, as reported in section 2 of this report, certain types of debt and assets tend not to be reported. This type of nonresponse might be important and should be studied further, but it will not be discussed here. An overview of item nonresponse can be found in the Variable Documentation VSB panel 1993 (Camphuis & Ketelaars, 1995).

### Biases caused by refusals to become panel members

Data are collected by using computers and modems and this requires that the respondents have some insight in how to use a computer, or that they are willing and able to learn it. This requirement might cause some biases. It is possible that elderly or people with little education and little knowledge about computers, would decline to participate in the survey. It is therefore possible that people co-operating in the telepanel are a little better educated than others. It is not unreasonable to expect that more highly educated people are more aware of alternative ways of holding assets than others, so it might happen that panel members engage in more complex saving behaviour than nonmembers.

According to Ritzema and Human (1991), previous investigations have shown that the procedure used to recruit the members to the telepanels minimizes the drop-out rate that can be attributed to the use of computers. The computers are only mentioned in the last part of the initial recruitment interviews. According to Ritzema and Human (1991) only 5.1% of those who chose not to participate in the INFO telepanel made this decision because a computer had to be used. If this is correct, sample bias caused by the use of computers might be small. We do, however, need more information about reasons for why some households do not wish to co-operate in the panel.

It is possible to adjust for biases caused by refusals in the recruitment process by using purposive selection or quota sampling. However, it is only possible to estimate and adjust for biases for variables of which we know the distribution in the target population, like income and age. Biases with respect to psychological traits are impossible to identify, and afterwards weighting or quota sampling cannot be done in order to secure representativeness for these variables. Neyman (1934) reports that Italian statisticians who applied the method of purposive selection did not find their results to be satisfactory. Their (Gini & Galvini, 1929) chief purpose had been to obtain a sample which would be representative of the Italian population with respect to important demographic, social, economic, and geographic characteristics. However, the comparison between the sample and the whole country showed disappointing results. In spite of the fact that the average values of the seven control variables used when constructing the sample were in satisfactory agreement, the agreement of average values of other variables, which were not used as controls, were poor. The selection of respondents with respect to socio-economic variables did not provide a sample which was representative with respect to other variables.

Calling the STP "the representative sample" might therefore not be justifiable. People co-operating in the panel might be different from the people refusing to do so. They might be more open, have a better organised economic behaviour or not feel that they have anything to hide. The low response rate can thus represent a threat to the extent in which results involving effects of psychological variables on economic behaviour can be generalised. The validity of studies of psychological factors effect on economic behaviour may thus be weak.

### Bias caused by questionnaire-specific nonresponse

There are many possible causes of questionnaire-specific nonresponse. As the panel members receive questionnaires rather frequently, it is probable that they sometimes do not answer the questions because they are on holiday or make trips in the weekends. This might cause a bias as we can expect the active and enterprising households to go away in weekends more often. People with higher income can afford weekend tour trips away from their homes more frequently. Alternatively, they might work more during weekends than others. Furthermore, it might also be the case that people only answer questionnaires they like. For example, people who refuse to

<table>
<thead>
<tr>
<th></th>
<th>Response rate within sample 1st wave</th>
<th>Overall response rate 1st wave</th>
<th>Length of membership in panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>STP</td>
<td>18.5 %</td>
<td>94.8 %</td>
<td>Up to 3 years</td>
</tr>
<tr>
<td>HIP</td>
<td>20.2 %</td>
<td>99.4 %</td>
<td>Up to 1 year</td>
</tr>
</tbody>
</table>

Table 3.3 Response Rates for the STP and the HIP for the 1st Wave

Footnote: More exact information about response rates is not available. The ratio given in table 3.3 is thus not actual response rates. As some of the households that initially participated in the panel dropped out before the data for the Savings Project were collected, the response rates in the columns \( A \) and \( A \times B \) should be lower.
answer the VSB-project questionnaires may do so either because they have large savings or little savings - the former do not wish to report the amount of their savings, while the latter might feel that they have little of interest to contribute to the study or may be reluctant to disclose how little savings they have. Panel members who perceive themselves as atypical with respect to financial affairs, might therefore drop out.

Table 3.3 shows that the questionnaire-specific response rate is higher for the HIP than for the STP. The reason for this can be that the members of this panel are all new panel members, and disclose how little savings they have. Panel members who perceive themselves as atypical with respect to financial affairs, might therefore be unmotivated to answer questions about financial data. HIP members also received better computers than the STP members, something that might increase their feeling of commitment to the Savings Project. The HIP members also get more time to answer the questionnaires, and this increases the probability of that they will be at home during the period the responses should be given.

3.5 Use of the Panels

Below, different kinds of samples are recommended for different kinds of tests and generalisations (Cook & Campbell, 1979). According to Calder et al. (1981) it is important to distinguish between theory research and effects research. In theory research the goal of the research is to expand the scope of existing scientific theories by testing whether they also apply to a specific population. In these types of studies observed effects in each study are used to assess the status of the theory. In effects research the goal of the research is to produce knowledge about specific situations and generalise this to other situations. It is the effects themselves that are generalised rather than the theoretical constructs and hypothesised theoretical network used for deducing patterns of outcomes.

When theory research is the goal of a study, Calder et al. (1981) recommend that the researcher should try to falsify the theory in focus by creating a research context which will provide a test of the theory as rigorous as possible. Internal validity should be given priority to external validity. Internal validity refers to which extent one can be sure of that the observed variation in a dependent variable is caused by the independent variable (Cook & Campbell, 1979). One implication of giving internal validity priority is that the sample used in the study should be as homogeneous as possible. This entails sampling from groups of individuals that are similar on dimensions likely to influence the variables of theoretical interest, but which is not a part of the theory in focus. Most scientific theories are universal in scope, and they should therefore apply to all kinds of respondent groups, regardless of whether they are representative of the population or not. A homogeneous sample permits more exact theoretical predictions than may be possible with a heterogeneous group. It also decreases the chance of making a false conclusion about whether there is covariation between the variables under study (Calder et al., 1981; Calder et al., 1982).

If the goal of the research is to study specific effects, Calder et al. (1981) recommend use of statistically representative samples. The reason for this is that effects research is based on the assumption of correspondence between the research and the real-world situation. Specific effects obtained in one study are expected to be similar to findings that would be observed if data were collected from other populations and situations. It is therefore important to secure that the real world phenomena in focus is represented in the study, and external validity should be given priority to internal validity. External validity refers to which degree results from one study can be applied to other situations and other populations beyond the sample (Cook & Campbell, 1979). Natural environments and natural behaviour should be used as independent and dependent variables respectively.

Following the advice from Calder et al. (1981), the best application of the data collected from the STP and the HIP is to use them for theory development and studies of specific effects. The STP is meant to be representative of the Dutch population, and the members of the panel are heterogeneous with respect to a large number of variables. Also the HIP is meant to be representative of high income households. Both panels are asked about natural behaviour in natural environments. The panels are therefore useful for testing special effects, as it might give an indication of whether a proposed relationship can be found in a real-world situation. However, due to the probability of that sample biases and panel conditioning might cause the samples to be atypical of the target populations, results should be replicated using different samples and different data collection methods. One should also be careful of using estimates derived from the panels as estimates for the target population.

3.6 Summary and Conclusions

Data from two different sub-panels are collected for the VSB-Center Savings Project. One sub-panel is intended to be representative of the Dutch population with respect to certain socio-economic variables. This panel was set up in 1991 by the Stichting TelePanel and is used for a variety of research projects. An EPSEM sampling scheme combined with quota sampling was used when recruiting respondents for the panel. The group sampling was applied to compensate for selective nonresponse in the recruitment process. Stratification was also applied in order to secure representation of all regions and larger cities. The annual attrition rate is 30% and continuous surveys are carried out in order to replace households that drop out with households with similar characteristics.

The other sub-panel, the HIP, consists of households from the upper 10% of the income distribution in the Netherlands. These households were recruited specifically for the VSB-Center Savings Project because of a wish to study the behaviour of high-income households in more detail. They now also serve as informants for other research projects, but not for as many as the members of the representative panel. The HIP members are quite homogeneous with respect to education, family composition, housing situation (owner or tenant) and other variables which typically are related to income.

The level of nonresponse is quite high for both samples (81-90% cf. table 3.3) and this might cause rather serious sampling biases. As shown in table 3.4, the nonresponse showed a certain
pattern, as old people and families without children were less willing to co-operate in the panel. This nonresponse bias was adjusted for by using quota sampling. The quota sampling, probably made the sample more representative with respect to the variables that were taken into consideration when selecting households. However, as pointed out by Neyman (1934), such a purposive selection of households does not necessarily provide a sample which is representative with respect to other variables. People who chose to become panel members might be different from those who refused to co-operate for example with respect to psychological or behavioural variables. These are variables for which we do not know the distribution in the population for, and adjustment for biases is therefore impossible. Such sample biases might be caused by the sampling frames applied and selective nonresponse. Researchers using the data should therefore carefully evaluate to which extent possible biases might influence their specific study.

4. Summary and Conclusions

The VSB-CentER Savings Project was started in 1990. The goals of the project are 1) to test the descriptive and predictive power of economic and psychological factors on households' saving behaviour and 2) to study the stability of any possible influence of these variables on saving behaviour. The fact that also effects of psychological factors are in focus, makes the data set collected for the VSB-CentER Savings Project unique. It contains indicators of a range of economic psychological concepts. This makes developing and testing of models of economic behaviour containing both economic and psychological determinants possible, as well as tests of interaction effects between these variables. This has only to a limited extent been possible before.

The data are collected by using a relatively new data collection method, Computer Assisted Panel Research (CAPAR). Members of the panels have agreed to answer a certain number of questions on a regular basis (once a week or once a month). This makes it possible to use such a long questionnaire as the one used for the Savings Project. The questionnaire can be split into separate items for which we do not know the distribution in the population for, and adjustment for biases is therefore impossible. Such sample biases might be caused by the sampling frames applied and selective nonresponse. Researchers using the data should therefore carefully evaluate to which extent possible biases might influence their specific study.

Also the respondents' evaluation of the questionnaire and their own responses is overall positive, and this also indicates that we might expect the data to be relatively accurate. As the response rates within the panels also are high (see table 5.3), it seems like the panel members are motivated to fulfill their obligations toward Schuitema TelePanel.

However, more research should be done in order to verify the effects of the panel membership. Panel members might change their behaviour due to their participation in the panel, and this might be a competing explanation of observed behaviour to other proposed causes, thus weakening internal validity of the studies. We need more knowledge about how and to what extent panel conditioning might influence the results from our analyses of the Savings Project data.

The sample used for the VSB-CentER savings project consists of two sub-samples. One sub-sample is intended to be representative of the Dutch population with respect to certain socio-economic variables. The households were selected using an EPS/SEM sampling design combined with quota sampling. The sampling design also involved some stratification in order to secure representation of all regions and larger cities. This panel (the STP) consists of individuals in about 2000 households. The other sub-sample is a result of stratified sampling, and is a sample of Dutch high-income households, consisting of individuals in about 1000 households (the HIP). The additional sample of high-income households gives us opportunities to test whether theories of household economic behaviour also applies to this particular group. Normally, a simple random sample does not provide enough high-income households to draw any conclusions about their behaviour.

One high level of nonresponse in the recruitment process for the panels, might also cause sample biases. The panels are representative with respect to certain socio-economic variables, as quota sampling was used in order to adjust for biases in the recruitment process. However, the quota sampling does not secure that the panels are representative with respect to other types of variables, like psychological and behavioural variables. People co-operating in the panel might be different from the people refusing to do so. They might be more open, have a better organised economic behaviour or not feel that they have anything to hide. Previous investigations also have shown that people that participate in savings studies have lower average savings than those who refuse to participate. The low response rate can thus represent a threat to the extent in which results involving effects of psychological variables on economic behaviour can be generalised.

In spite of the weaknesses related to the data collection method and the sampling procedures discussed in this report, the data collected for the Savings Project represent a unique possibility with respect to developing new theories of households' saving behaviour. Depending on how
many waves of data collection that will be carried out, also the second goal of the Savings Project can be fulfilled; testing the stability of any possible influence of economic and psychological variables on saving behaviour. The weaknesses described in this report will not influence every result from analyses of the data to the same degree. However, researchers using the data should be aware of their possible effects, and take them into account when evaluating results and before generalising results to other populations and situations.

Reference List
